ALMA MATER STUDIORUM - UNIVERSITA' DI BOLOGNA

ARCES - ADVANCED RESEARCH CENTER ON ELECTRONIC SYSTEMS
FOR INFORMATION AND COMMUNICATION TECHNOLOGIES E. DE CASTRO

SCUOLA DI DOTTORATO IN TECNOLOGIE DELL'INFORMAZIONE

CICLO XXI - ING-INF/01

MODELLING AND SIMULATIONS OF POST–CMOS DEVICES

Tesi di Dottorato

Presentata da Stefano Poli

Coordinatore della Scuola di Dottorato:

Chiar.mo Prof. Ing. CLAUDIO FIEGNA

Relatore:

Chiar.mo Prof. Ing. GIORGIO BACCARANI

Esame Finale: anno 2009

Contents

Contents i

List of symbols and abbreviations iii

Abstract ix

Sommario xi

Introduction xiii

I QUANTUM TRANSPORT SIMULATION 1

1 Quantum transport analysis: the NEGF formalism 3
- 1.1 Poisson-Schrödinger simulations 4
- 1.2 The NEGF formalism: a phenomenological introduction 5
- 1.3 Resolution on a discrete domain 9
  - 1.3.1 Real-Space approach 9
  - 1.3.2 Mode-Space approach 11
- 1.4 The electron-phonon interaction 15
  - 1.4.1 Phonon scattering in silicon nanowires 17
  - 1.4.2 Phonon scattering in carbon nanotubes 20
- 1.5 The Poisson equation 21
- 1.6 Summary 23

II APPLICATION TO THE FETs ANALYSIS 29

2 Transport analysis in CNT based FETs 31
- 2.1 Electronic properties of carbon nanotubes 32
- 2.2 Device architecture 37
- 2.3 Non-parabolicity and band-to-band tunneling 38

2.4 Performance evaluation of CNT-FETs 46
    2.4.1 Fixed gate length 46
    2.4.2 Effect of gate length scaling 48
    2.4.3 Effect of changing gate dielectric 48
    2.4.4 Effect of CNT chirality 49
    2.4.5 CNT-FET with underlapped source/drain regions 51
2.5 Performance evaluation of CNT Tunneling-FETs 53
    2.5.1 Fundamental transport properties 54
    2.5.2 Oxide thickness and voltage variations 57
    2.5.3 Scaling trends 60
2.6 Summary 62

3 Transport analysis in Si-NW based FETs 67
3.1 Electronic properties of SiNWs 68
3.2 Device Model 70
3.3 Size dependence of SR–limited mobility 73
    3.3.1 SR generation 74
    3.3.2 Effective mobility extraction 75
    3.3.3 Ballistic mobility analysis 77
    3.3.4 Impact of surface roughness 79
3.4 Impact of remote-Coulomb scattering 82
    3.4.1 Fixed charges generation 83
    3.4.2 Screening effect 85
    3.4.3 Transfer characteristics analysis 86
    3.4.4 RCS–limited mobility 89
3.5 Dissipative transport: the impact on device scaling 92
    3.5.1 Transfer characteristic analysis 92
    3.5.2 Short-channel effective mobility 95
3.6 Summary 98

Conclusion 107

Curriculum Vitae 109

List of symbols and abbreviations

Symbols

a	Bohr radius
a_1, a_2	Graphene lattice characteristic vectors
a_{cc}	Carbon-to-carbon bonding distance
a_q	Phonon annihilation operator
${a_q}^\dagger$	Phonon creation operator
$c_i^{n,m}$	Correlation term
$\mathbf{C}_h$	Chiral vector
d_{CNT}	Carbon nanotube diameter
$D^<$	Phonon less-than Green's function
$D^>$	Phonon greater-than Green's function
δ	Kronecker delta
	Dirac delta
Δ	Discretization step
ϵ	Dielectric constant of the medium
ϵ_{Si}	Silicon dielectric constant
ε	One-dimensional subband profile
E	Energy
$E(\mathbf{k})$	Energy dispersion relation
E_c	Conduction band-edge
E_g	Band gap
E_i	Neutrality level
E_v	Valence band-edge

f	Fermi function
$\mathcal{F}_\eta$	η-order Fermi-Dirac integral
g^r	Surface Green's function
G	Conductance
$G^<$	Less-than Green's function
$G^>$	Greater-than Green's function
G^r	Retarded Green's function
Γ	Gamma function
$\hbar$	Reduced Plank's constant
H	Hamiltonian operator
H_d	Device Hamiltonian operator
I	Current
	Identity matrix
$I_{m,\nu'}^{n,\nu}$	Form factor
$\mathbf{J}$	Current density vector
J_x, J_y, J_z	Current density vector components
$\mathbf{k}$	Lattice wave vector
k_x, k_y, k_z	Lattice wave vector components
k_b	Boltzmann constant
$\mathbf{K}_j$	Momentum transfer
$\tilde{\mathbf{K}}_a$	Carbon Nanotube electron-phonon matrix element
Λ	Density spectral function
$\mathbf{l}$	Graphene translation vector
n	Electron density
$m^\star$	Effective mass
m_0	Free electron mass
m_l	Silicon longitudinal effective mass
m_t	Silicon transverse effective mass
M_q	Silicon electron-phonon matrix element
μ	Chemical potential
μ_{bal}	Ballistic mobility
μ_{eff}	Effective mobility
μ_{ph}	Phonon limited mobility
μ_{RCS}	Remote-Coulomb scattering limited mobility
μ_{sc}	Scattering limited mobility
μ_{SR}	Surface-roughness limited mobility

N	Net doping concentration
N_c	Silicon effective density of states
N_q	Average phonon density
N_D	Drain doping concentration
N_S	Source doping concentration
ω_q	Phonon angular frequency
Ω	Box volume
p	Hole density
ϕ	Electrostatic potential
	Two-dimensional wave function
ψ	Wave function
q	Electron elementary charge
	Phonon wave vector modulus
r	Linear transport regime backscattering coefficient
Q	Total linear charge
$\mathbf{q}$	Phonon wave vector
$\mathbf{r}$	Position vector
ρ	Charge density
	Silicon density
ρ_{1D}	One-dimensional carbon nanotube mass density
Σ_{lc}	Left contact self-energy
Σ_{rc}	Right contact self-energy
Σ_{ph}	Phonon self-energy
Σ^r	Retarded self-energy
$\Sigma^<$	Less-than self-energy
$\Sigma^>$	Greater-than self-energy
σ_b	Subband degeneracy
σ_s	Spin degeneracy
σ_v	Valley degeneracy
t_x	One-dimensional transfer integral
τ	Delay time
T	Transmission probability
T	Lattice temperature
$\mathbf{T}$	Carbon nanotube translational vector
u_l	Silicon sound velocity
U	Potential energy

v_a	Carbon nanotube sound velocity
$v_{\mathrm{pp}\pi}$	π-bonding energy
V	Base change matrix
V_{xc}	Exchange-correlation potential
Ξ	Deformation potential

Abbreviations

1–D	One-dimensional
2–D	Two-dimensional
3–D	Three-dimensional
BTBT	Band-to-band tunneling
CEM	Constant Effective Mass
CMOS	Complementary Metal-oxide-semiconductor
CMS	Coupled Mode Space
CNT	Carbon Nanotube
DFT	Density Functional Theory
DG	Double-gate
DIBL	Drain induced barrier lowering
EM	Effective mass
FET	Field Effect Transistor
GAA	Gate-All-Around
HIBL	Hole induced barrier lowering
HP	High Performance
$\mathrm{I_{bal}}$	Ballistic current
$\mathrm{I_{scat}}$	Scattering-limited current
$\mathrm{I_{ON}}$	On-current
$\mathrm{I_{OFF}}$	Off-current
MOS	Metal-oxide-semiconductor
NEGF	Non-equilibrium Green's Function
LOP	Low Operating Power
PSD	Power density spectrum
RMS	Root Mean Square
RS	Real Space
SS	Inverse sub-thereshold slope
NW	Nanowire

TB	Tight-Binding
UMS	Uncoupled Mode Space
VEM	Variable Effective Mass
ZF	Zone Folding

Abstract

As the dimensions of field-effect-transistors (FETs) scale down to the deca-nanometric range, semi-classical models start to fail in describing fundamental transport properties. In addition, as new channel materials are introduced some interesting quantum phenomena, usually neglected in the analysis of silicon-based devices, grow in importance. A full quantum description is thus required to have valuable predictions for innovative FETs.

In this framework, the Non Equilibrium Green's Formalism has been successfully used for a full-quantum description of transport in nanostructures within different approximations and models. Here, it is presented for the realization of a numerical solver for self-consistent Poisson-Schrödinger simulations within the effective-mass approximation. Both ballistic and dissipative transport are considered. The inclusion of the electron-phonon interaction is presented in the self-consistent Born approximation.

As a first alternative to standard silicon metal-oxide-semiconductor FETs (MOSFETs), carbon nanotube based FETs are presented. Due to the interesting electronic properties of this material and the valuable expectations raised from the first experimental FETs realizations, carbon nanotubes as novel transistors channel material have received large interest. In the used transport model, corrections to the standard parabolic effective-mass approximation are presented for the inclusion of the band-to-band tunneling and to account for the non-parabolicity of the one-dimensional (1-D) energy dispersion relation. Two different device architectures are presented showing a detailed analysis of the impact of electron-phonon interaction and of the limiting transport factors as architectural and geometrical characteristics of the device are changed.

Not only new channel materials but also new architectural concepts can be exploited to extend the scaling process while keeping increasing devices performance. For this reason, silicon-nanowire FETs are among the most promising 1-D structures to achieve the ballistic transport limit. Fabrication defects are in any case important to be included. Here, the presence of rough silicon-silicon oxide interfaces and of fixed charge centers in high-κ materials gate stacks is considered as a scattering source additional to the electron-phonon interaction. Surface-roughness and remote-Coulomb interaction are phase coherent scattering and treated via a direct definition of the perturbations for every single device, accordingly to a given statistical distribution. The variability of performance is thus studied on a set of sample devices mainly focusing on the extension of the low-field mobility concept to short-channel silicon-nanowire FETs using a full-quantum description for all considered scattering mechanisms in a three-dimensional (3-D) solver.

Sommario

Con la riduzione delle dimensioni dei transistori ad effetto di campo (*fieled-effect-transistor*, FET) fino a scale decananometriche, modelli semi-classici per il trasporto cominciano a fallire nel descrivere alcune proprietá fondamentali. In aggiunta, con l'introduzione di nuovi materiali di canale alcuni interessanti fenomeni quantistici, usualmente non considerati per dispositivi in silicio, crescono in importanza.

In questa ottica, il formalismo delle funzioni di Green di fuori equilibrio (*Non Equilibrium Green's Function*) stato utilizzato con successo per la descrizione completamente quantistica del trasporto in nanostrutture secondo diverse approssimazioni e modelli. In questo lavoro, viene introdotto per la realizzazione di un solutore numerico per simulazioni Schrödinger-Poisson autoconsistenti secondo l'approssimazione di massa efficace. Sia il trasporto di tipo balistico che dissipativo sono considerati. L'inclusione dell'interazione elettrone-fonone presentata nella approssimazione di Born autoconsistente.

Una prima alternativa ai dispositivi MOSFET in silicio é rappresentata dai transistor ad effetto di campo basati su Nanotubi di Carbonio (*Carbon Nanotube*). Grazie alle interessanti proprietá elettroniche di questo materiale e alle validi aspettative derivanti dalle prime realizzazioni sperimentali di FET, l'impiego dei nanotubi di carbonio come nuovo materiale di canale per transistori ha attratto largo interesse. Per il modello di trasporto utilizzato, sono presentate correzioni alla approssimazione di massa efficace parabolica al fine di includere la descrizione del tunneling banda a banda (*band-to-band tunneling*) e per tenere in conto della non parabolicit della relazione di dispersione mono-dimensionale. Sono quindi presentate due differenti architetture del dispositivo, mostrando una analisi dettagliata dell'impatto dell'interazione elettrone-fonone e dei fattori limitanti per il trasporto a fronte di variazioni architetturali e geometriche del dispositivo.

Non solo nuovi materiali di canale, ma anche nuove archietture possono essere sfruttate per estendere il processo di *scaling* garantendo un aumento delle prestazioni dei dispositivi. Per questo motivo, i transistor a nanofilo di silicio (*silicon Nanowire*) sono considerati tra le strutture mono-dimensionali pi promettenti

al fine del raggiungimento di una condizione di trasporto balistico. La presenza di difetti di fabbricazione deve essere per tenuta in conto. Nel presente lavoro, interfacce rugose tra silicio ed ossido di silicio sono prese in considerazione cosí come la presenza di centri di cariche fisse quando ossidi di *gate* compositi con materali ad elevata costante dielettrica (*high-*κ) sono introdotti. Entrambi i difetti costituiscono delle sorgenti di *scattering* addizionali alla interazione elettrone-fonone. La rugositá d'interfaccia e l'interazione Coulombiana remota sono collisioni mantenti la coerenza di fase e sono trattati mediante una diretta definizione delle perturbazioni per ogni singlo dispositivo considerato, in accordo ad una data distribuzione statistica. La variabilitá delle prestazioni studiata quindi su di un insieme di dispositivi campione focalizzandosi primariamente sull'estensione del concetto di mobilitá di bassi campi a FET di nanofilo di silicio a canale corto utilizzando un descrizione completamente quantistica per i meccanismi di collisione cosinderati in un solutore tri-dimensionale.

Introduction

The request for increasing performance and reduced area occupancy for both active and passive electronic components pushes the scaling process every day closer to the physical limits of the silicon-based MOS technology. Therefore, new type of nanoscale devices are being widely investigated. The use of both new architectures and new channel materials can overcome the limitations that nowadays nanoscale MOSFETs experience, i.e. short channel effects and gate leakage.

Since their discovery, carbon nanotubes (CNTs) have received large attention. The possible application for nanoelectronic devices has been extensively explored since the demonstration of the first carbon nanotube transistors and rapid advancements have been achieved in the understanding of the physical properties. Nearly ballistic transport in low bias regime, compatibility with high-κ dielectrics due the absence of superficial dangling bonds are promising features in electronic applications. In addition, due to the direct bang gap band structure, enhanced optical emission can lead to optoelectronic applications, while, for metallic nanotubes, the high resistance to electromigration is promising for the use in interconnections.

At the same time, semiconductor nanowires (NWs) are attractive building blocks for the future electronic applications. Numerous examples of nanowire fabrication and integration have been recently reported through the use of different channel materials (e.g. Si, Ge, GaAs) and lateral dimensions and shapes. Among the various possibility, silicon is still reported as the leading investigation material, due to the compatibility with the standard electronic industry and being silicon nanowires a natural extension of the scaling process of MOSFET devices. In particular, the use of a *top-down* fabrication approach, combining advanced chemical and lithographic process, has proved its technological compatibility with the standard complementary metal-oxide-semiconductor (CMOS) process through the integration of both single devices and logic circuits. On the other hand, the use of a *bottom-up* approach gives the possibility of synthesizing nanowires in single-crystalline form with precisely controlled diameter, length and chemical composition. The composition of new exotic core/shell heterostructures has also been reported. In addition, the *bottom-up* fabrication does not suffer for high fab investments allowing academic structures to keep on contributing to the growth of nanowire electronics.

In the process of development of novel technologies and designs, the use of technology computer aided design is of fundamental importance for reducing time and investments costs. Many are the possible levels of abstraction going down from compact models for circuit simulations to more physics based models for the analysis of transport in single devices. In the presented work, the aim has been developing simulation tools for the design and study of post-CMOS devices, focusing on carbon nanotube and silicon nanowire based field-effect-transistor, relying on the state-of-the-art of semiconductor physics.

The necessity for a proper description of transport on nanometric scale has pushed to abandon some classical concepts and look with greater interest to quantum mechanical models. Quantum confinement with its effect on band structure and carrier distribution or tunneling effect are among the most important quantum mechanical phenomena to be considered in order to achieve a valuable prediction on device performance.

When developing a simulation tool, importance has to be given to the trade-off between the accuracy in the physical description and the required computational burden. The use of the Non Equilibrium Green's function (NEGF) formalism has been considered for the analysis of the transport problem, giving the tools for a description of both ballistic and dissipative transport on different levels of approximation for the electronic structure. In recent years, both semi empirical tight-binding (TB) and effective-mass (EM) models have been used in simulations of quantum transport in SiNW- and CNT-FETs.

The tight-binding model allows for the description and the inclusion of the entire band structure of the material by means of a variable number of parameters, i.e. the number of atomic orbitals considered. Satisfactory results can be obtain with a single orbital for carbon based material, while requiring 10 or 20 different orbitals for silicon. The computational cost is highly affected by the number of orbitals as well as by the dimension of the device, i.e. by the total number of atoms to be accounted for.

On the other side, at the price of an approximated band structure description, the effective-mass approach aims at describing quantum transport by means of two parameters, i.e. the band gap and the effective mass, allowing for a drastic reduction of computational costs. Larger devices can thus be analyzed with less dependence of the computational time on the device dimensions. Some of the limitations of the EM approximation can be removed by the use of correction for the inclusion of non-parabolicity or of the band-to-band tunneling (BTBT) effect. In addition, although nanoscaled transistors operate closer and closer to the ballistic regime, the inclusion of the electron-phonon interaction is still a fundamental aspect in the analysis of the device. Phonon-scattering can thus be introduced in a full-quantum description with reasonable computational costs, otherwise non affordable in a TB description.

The presence of other mechanisms limiting the transport properties has also to be considered. Defects introduced in the fabrication process affecting the device structure and inducing additional scattering process are not negligible, in particular during the early stage of a device development. The process variability has recently become an important field of research intimately connecting the world of numerical simulations with that of device fabrication.

Therefore, the efforts have been focused on the realization of a full-quantum simulator, based on the self-consistent solution of the Schrödinger and Poisson equation within the NEGF formalism and the effective-mass approximation. The inclusion of the electron-phonon interaction is considered as well as, in the analysis of SiNW-FETs, the effect of surface-roughness and remote-Coulomb scattering.

In Chapter 1 a brief overview of the NEGF formalism is given. The treatment is focused on the application of the formalism to self-consistent device simulations, describing the fundamental equations for the evaluation of the physical observables of interest (i.e. carrier density and current density) in steady-state condition. The Coupled Mode Space approach, for the simulation of transport on the space of coordinates is presented in the EM approximation with the derivation of proper boundary conditions and of the electron-phonon interactions.

Chapter 2 is focused on the application of the described formalism to the simulation of CNT-based FETs. A description of the fundamental electronic properties of CNTs is presented as obtained through a simple single-orbital TB model. Corrections to the parabolic effective-mass approximation are included to account for a better description of vanishing states and BTBT with a direct comparison with TB transport simulations. Two different device architectures have been analyzed showing the fundamental transport mechanisms, the important role played by the electron-phonon collisions and thus providing an investigation of the ultimate scaling possibilities for CNT-FETs.

In Chapter 3 the analysis is dedicated to SiNW-FETs. A parabolic effective-mass model is used focusing on wire with lateral dimensions superior to 3 nm in a linear transport regime. The focus is more on the analysis methodology, with an investigation of the impact of different scattering mechanisms on short-channel devices effective-mobility through the use of a statistical approach. Surface roughness, remote-Coulomb scattering and the electron-phonon interaction are considered, giving a complete tool complementary to semi-classical models.

Part I

QUANTUM TRANSPORT SIMULATION

Chapter 1

Quantum transport analysis: the NEGF formalism

As the dimensions of electronic devices scales down to the deca-nanometric regime and the choice for new channel materials and for new architectures is adopted, a new puzzling task has raised: a multi-scale physics approach to the device simulation. In this framework, methods moving from first-principle approaches (DFT, *Ab initio*) to the effective-mass approximation should be considered focusing on different levels of abstraction. The latter is more suitable for studies on the optimization of the device, the former are important for the calibration of the parameters used in higher level approximations.

At the same time, the increasing importance of physical phenomena like quantum tunneling, reflections through metallic/semiconducting barriers and through channel potential barriers, strong quantum confinement and band-to-band tunneling (BTBT) have forced to look with much more interest at a full-quantum description of carrier transport. Hence, device numerical simulations based on the self-consistent solution of the Schrödinger and Poisson equations are envisaged in order to correctly model these physical effects and give a valuable analysis and prediction of the device performance.

The Non Equilibrium Green's function formalism is being more and more used in the numerical simulation of transport for low-dimensional systems. A rigorously quantum mechanical approach, independent of the chosen basis used in the description, as well as the possibility of the treatment of the electron-phonon interaction, have been soon highlighted as strong points of this theory.

This chapter is meant to give a general overview of the physical models and numerical methods used in this work for the quantum transport analysis. The fundamental equations and basic concept of self-consistent simulation are given Sec. 1.1. A phenomenological introduction to the NEGF formalism is presented

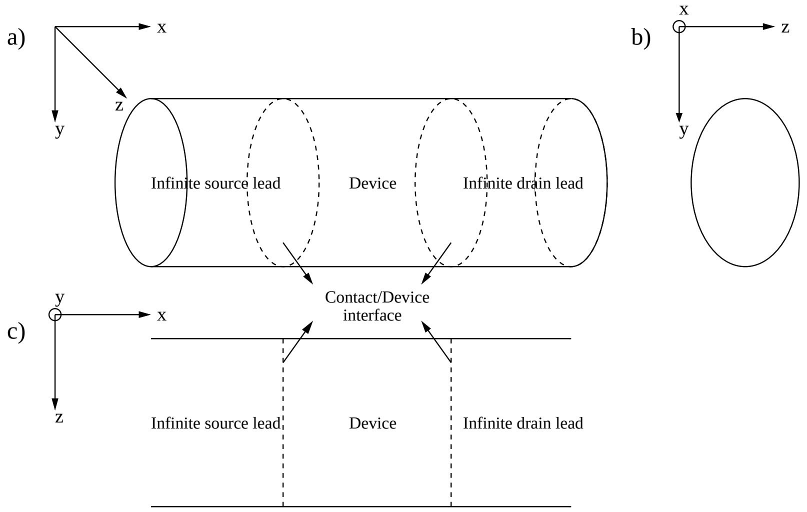

Figure 1.1: *Schematic representation of a generic reference device. a) Three-dimensional view; b) Cross-section on transverse plane (confinement direction); c) Cross-section along the longitudinal axis (transport direction).*

in Sec. 1.2, while two possible approaches for the solution of the fundamental kinetic equation, namely the Coupled Mode Space and uncoupled mode space, are derived in Sec. 1.3. Finally, the framework for the introduction of the electron-phonon interaction is presented in Sec. 1.4.

1.1 Poisson-Schrödinger simulations

A sketch of the reference device structure for the numerical simulations is given in Fig. 1.1, representing a 3-D view and both longitudinal and transverse two-dimensional (2-D) sections. The complete domain is supposed to be formed by a device region with a finite extension and two contact regions infinite along the transport direction x (section c) of Fig. 1.1). The domain has a finite width on the transverse plane. Due to the double confinement on the yz-plane (section b) of Fig. 1.1), only one direction of free propagation is left, identifying the device as a quasi-1D system. Although no restrictions are present on the geometry definition, in the following of the chapter, a rectangular shaped domain and a rectangular discretization mesh will be supposed in deriving the equations, allowing for clearer analytical formulations and considerations on the computational costs.

The calculation of the fundamental physical observables of the system (i.e. charge and current distribution) is obtained, in steady-state conditions, by the solution of the time-independent Schrödinger equation, that in the parabolic effective-mass approximation for a single electron in a 3-D domain reads out

$$\left(-\frac{\hbar^2}{2}\nabla\cdot\frac{1}{m^\star(x,y,z)}\nabla\ +U(x,y,z)\right)\psi(x,y,z)=E\psi(x,y,z)\ , \tag{1.1}$$

where ψ is the electron wave function, $m^\star$ is the position-dependent effective mass tensor, U the conduction band potential energy profile, and E the injection energy of the carrier. The presence of an external or induced magnetic field is neglected. Equation (1.1) has to be solved with proper boundary conditions depending on the chosen geometry of the device and depending on the model adopted for the source and drain contacts.

The electrostatic potential ϕ is computed through the solution of the Poisson equation

$$-\nabla\cdot(\epsilon(\mathbf{r})\nabla\phi(\mathbf{r}))=\rho(\mathbf{r}) \tag{1.2}$$

where $\epsilon(\mathbf{r})$ is the position-dependent dielectric constant of the medium, ρ is the charge density, and $\mathbf{r}=(x,y,z)$ is the 3-D position vector. The charge density is given by the sum of the contributions due to free carriers, doping concentration and other fixed charge centers, as for example charges trapped in the oxide. The solution depends on the boundary condition (e.g. electrostatic potential at the electrodes) and, if included, it accounts for the additional exchange-correlation term (see Chapter 3).

The electrostatic potential enters in (1.1) through the determination of the conduction band profile ($U \leftrightarrow -q\phi$, where q is the elementary charge). On the other hand, the electron density depends on the solution of the Schrödinger equation through the square modulus of the electron wave function. Thus, a coupling between (1.1) and (1.2) is determined and an iterative self-consistent scheme has to be applied for the determination of the solution. As reported in the diagram of Fig. 1.2, starting from an initial guess for the solution of the Poisson equation, the two equations are iteratively solved till a convergence criteria is verified. The calculation of all physical quantities of interest will be performed on the last obtained solution. The specific methods used for the calculation of the physical quantities of interest will be described in the following Sections with the introduction of the solution of the transport problem with open boundary condition in the space of transverse modes (or Mode Space approach).

1.2 The NEGF formalism: a phenomenological introduction

The Non-equilibrium Green's function formalism provides a microscopic theory for quantum transport. If in the case of balistic transport the NEGF formalism is equivalent to the Landauer-Büttiker description of mesoscopic phenomena [1, 2], the real power of NEGF is to provide a general approach for describing dissipative transport, combining quantum dynamics and a statistical description of the interactions [3].

A meaningful and complete treatment of the NEGF formalism requires to refer to the quantum field theory and it is out of the aim of the thesis. Several

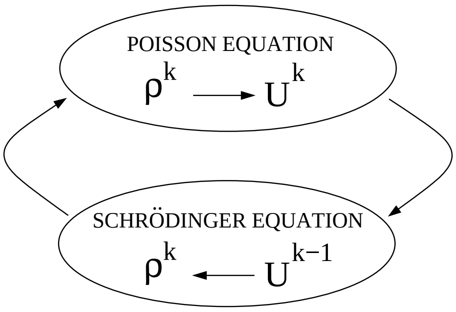

Figure 1.2: *Diagram representing the iterative self-consistent solution of the Schrödinger and Poisson equations. Convergence criteria can be set on the maximum variation of the potential between two subsequent iterations.*

approaches are possible and various books have been devoted to the topic, e.g. [4, 5]. In the following, a steady state analysis is considered. In this framework, the Fourier transform of all the operators, and hence of all physical quantities, of interest allows for a description in terms of space and energy coordinates instead of space and time [3]. A matrix notation with respect to the discrete spatial coordinates will be adopted.

The kinetic equations describing the non-equilibrium transport in the system, as derived from Dayson's equations [4], according to the previous assumptions, can be expressed as [3]:

$$\begin{aligned} &[(E+i\eta)I - H_d - \Sigma^r]\, G^r = I \\ &G^{<,>} = G^r \Sigma^{<,>} G^{r\dagger} \ . \end{aligned} \tag{1.3}$$

In the first equation of (1.3), G^r represents the retarded Green's function, H_d the effective-mass Hamiltonian of the isolated device (left hand side operator of (1.1)), Σ^r, named retarded self-energy, represents the interactions of the device with external systems, E is the energy of the particle, η is a positive infinitesimal quantity, and † represents the Hermitian-transpose operator. In the second equation of (1.3) the less-than and greater-than functions, $G^<$ and $G^>$, describe the hole and electron density spectra, and the less-than and greater-than self-energy, $\Sigma^<$ and $\Sigma^>$, are related to the in-scattering and out-scattering probabilities to the device from the interacting systems. This set of equations gives the information about the dynamical and statistical properties of the system. In presence of the electron-phonon interaction, the two equations are coupled through the definition of the retarded self-energy (see Sec. 1.4).

The retarded self-energy results from the sum of three terms:

$$\Sigma^r = \Sigma^r_{\mathrm{lc}} + \Sigma^r_{\mathrm{rc}} + \Sigma^r_{\mathrm{ph}} \ , \tag{1.4}$$

representing the interaction with the left contact (Σ^r_{lc}), the right contact (Σ^r_{rc}), and with the phonon bath (Σ^r_{ph}). The explicit dependence of each term on the energy has been omitted. In the following, the left and right contacts will be modeled as two reservoirs at thermal equilibrium with the chemical potential μ_s and μ_d, respectively, infinite along the transport direction, replicating uniformly the same confining potential $U(y, z)$, and reflectionless. This model is used to mimic the inclusion of ideally ohmic contacts. From a general point of view, the contact self-energy functions allow for the inclusion of the effect of inward and outward waves propagating through the device from an infinitely extended region maintaining a representation on a finite discrete space. Their definition requires the evaluation of the retarded Green's function associated to the semi-infinite portion of contacts outside the device and can be done analytically or numerically depending on the used Hamiltonian representation (details in Sec. 1.3).

The inclusion of the electron-phonon interaction and additional details for the evaluation of the phonon self-energy will be presented in Sec. 1.4. The inclusion of this collision mechanism will be approached within a low-order approximation (self-consistent Born approximation) [4], able, in any case, to capture all the fundamental properties of the electron-phonon interaction in a full-quantum description.

The electron density and the three components of the current density are determined as the expectation value of the respective field operator and they can be expressed, for a given energy E as:

$$n(\mathbf{r}; E) = -\frac{i}{2\pi} G^{<}(\mathbf{r}, \mathbf{r}; E) \; , \tag{1.5}$$

$$\mathbf{J}(\mathbf{r}; E) = \pm \frac{q\hbar^2}{2m^\star} \lim_{\mathbf{r}' \to \mathbf{r}} (\nabla' - \nabla) G^{<}(\mathbf{r}, \mathbf{r}'; E) \tag{1.6}$$

where (1.6) is valid for single-band or multi-band band-structure models [6] and it will be specialized in the following section with explicit reference to the real-space and mode-space representations.

The presented description can be repeated also for the transport treatment in terms of holes. It can be demonstrated that the hole density at a given energy reads out [3]

$$p(\mathbf{r}; E) = \frac{i}{2\pi} G^{>}(\mathbf{r}, \mathbf{r}; E) \; , \tag{1.7}$$

where the less-than and the greater-than Green's functions verify the relation

$$i\left(G^{>} - G^{<}\right) = i\left(G^{r} - G^{r\dagger}\right) = \Lambda \tag{1.8}$$

being Λ the density spectral function, describing the total available electronics states, filled and not, at a given energy.

The imposition of an equilibrium condition for the carrier distribution in the contacts is fulfilled through the definition of the less-than and greater-than self-energy functions. They are expressed by the following relations [3]

$$-\,i\,\Sigma^{<}_{\alpha}(E) = \Gamma_\alpha(E) f(E - \mu_\alpha) \quad i\,\Sigma^{>}_{\alpha}(E) = \Gamma_\alpha(E)(1 - f(E - \mu_\alpha)) \tag{1.9}$$

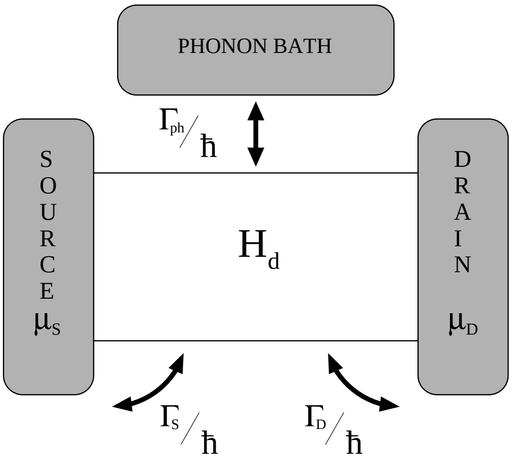

Figure 1.3: *Schematic representation of the interaction model between the device and the outer systems (contacts and phonon bath). The frequency of the interactions is mathematically expressed by proper* Γ *functions.*

for $\alpha = \mathrm{lc}, \mathrm{rc}$, and where $f(x) = 1/(1 + \exp(x/k_b T))$ is the Fermi-Dirac distribution, k_b is the Boltzmann constant, T is the temperature, and

$$\begin{aligned} \Gamma_\alpha(E) &= i\left[\Sigma^r_\alpha(E) - \Sigma^{r\,\dagger}_\alpha(E)\right] \\ &= i\left[\Sigma^>_\alpha(E) - \Sigma^<_\alpha(E)\right] \ . \end{aligned} \tag{1.10}$$

The physical meaning of the Γ functions defined by (1.10) allows for a better understanding of the self-energy description of scattering mechanisms. The quantity $1/\tau = -i\Gamma/\hbar$ represents a scattering probability describing the decay-rate of a carrier from an external system to the device and vice-versa. In the case of an ideal ballistic conductor, the in- and out- scattering from the contacts remains the only interaction mechanism affecting the propagation of the carrier along the device. This physical description holds for both left and right contacts and for self-energy functions representing the electron-phonon interaction. In the latter case, as described in Fig. 1.3, the scattering rate gives the probability of interaction of the electron in the device with the outer system of phonons, approximated as an unperturbed phonon bath. Interactions with the system are possible at every point of the device along the transport direction. In this general picture, $1/\tau$ describes the total phase relaxation rate [7].

As a consequence of the presented relations, in the case of an ideal ballistic conductor it is more evident the division between the dynamical and the statistical description of the systems. Introducing the formulation of the contacts less-than Green's functions in (1.3) and making use of (1.8), the electron density at a fixed energy and device position is given by

$$2\pi n(\mathbf{r}; E) = \Lambda_s(\mathbf{r}; E) f_s + \Lambda_d(\mathbf{r}; E) f_d \ , \tag{1.11}$$

where Λ_s and Λ_d are the density spectral function due to the injection from the source and drain, respectively. The contact equilibrium statistics f_s and f_d give the appropriate filling of the density of states. The knowledge of the retarded Green's function is required in the calculation of Λ independently of the knowledge of its occupancy, thus uncoupling the kinetic equations.

1.3 Resolution on a discrete domain

1.3.1 Real-Space approach

A description in the effective-mass approximation in the space of real coordinates has been supposed so far. In this picture, the discrete Hamiltonian operator H_d acting on the wave function component $\psi_{i,j,k} = \psi(x = x_i, y = y_j, z = z_k)$, as derived from the box-integration-method, is expressed by

$$
\begin{aligned}
&H_d\psi|_{i,j,k} = U_{i,j,k}\Omega_{i,j,k} - \\
&-\frac{\hbar^2}{8}\Bigg\{\frac{\psi_{i+1,j,k}-\psi_{i,j,k}}{\Delta_{x_i}}\left[\frac{\Delta_{y_{j-1}}\Delta_{z_k}}{m_x^{(i,j-1,k)}}+\frac{\Delta_{y_j}\Delta_{z_k}}{m_x^{(i,j,k)}}+\frac{\Delta_{y_{j-1}}\Delta_{z_{k-1}}}{m_x^{(i,j-1,k-1)}}+\frac{\Delta_{y_j}\Delta_{z_{k-1}}}{m_x^{(i,j,k-1)}}\right]+ \\
&+\frac{\psi_{i-1,j,k}-\psi_{i,j,k}}{\Delta_{x_{i-1}}}\left[\frac{\Delta_{y_{j-1}}\Delta_{z_k}}{m_x^{(i-1,j-1,k)}}+\frac{\Delta_{y_j}\Delta_{z_k}}{m_x^{(i-1,j,k)}}+\frac{\Delta_{y_{j-1}}\Delta_{z_{k-1}}}{m_x^{(i-1,j-1,k-1)}}+\frac{\Delta_{y_j}\Delta_{z_{k-1}}}{m_x^{(i-1,j,k-1)}}\right]+ \\
&+\frac{\psi_{i,j+1,k}-\psi_{i,j,k}}{\Delta_{y_j}}\left[\frac{\Delta_{x_i}\Delta_{z_k}}{m_y^{(i,j,k)}}+\frac{\Delta_{x_{i-1}}\Delta_{z_k}}{m_y^{(i-1,j,k)}}+\frac{\Delta_{x_i}\Delta_{z_{k-1}}}{m_y^{(i,j,k-1)}}+\frac{\Delta_{x_{i-1}}\Delta_{z_{k-1}}}{m_y^{(i-1,j,k-1)}}\right]+ \\
&+\frac{\psi_{i,j-1,k}-\psi_{i,j,k}}{\Delta_{y_{j-1}}}\left[\frac{\Delta_{x_i}\Delta_{z_k}}{m_y^{(i,j-1,k)}}+\frac{\Delta_{x_{i-1}}\Delta_{z_k}}{m_y^{(i-1,j-1,k)}}+\frac{\Delta_{x_i}\Delta_{z_{k-1}}}{m_y^{(i,j-1,k-1)}}+\frac{\Delta_{x_{i-1}}\Delta_{z_{k-1}}}{m_y^{(i-1,j-1,k-1)}}\right]+ \\
&+\frac{\psi_{i,j,k+1}-\psi_{i,j,k}}{\Delta_{z_k}}\left[\frac{\Delta_{x_i}\Delta_{y_j}}{m_z^{(i,j,k)}}+\frac{\Delta_{x_{i-1}}\Delta_{y_j}}{m_z^{(i-1,j,k)}}+\frac{\Delta_{x_i}\Delta_{y_{j-1}}}{m_z^{(i,j-1,k)}}+\frac{\Delta_{x_{i-1}}\Delta_{y_{j-1}}}{m_z^{(i-1,j-1,k)}}\right]+ \\
&+\frac{\psi_{i,j,k-1}-\psi_{i,j,k}}{\Delta_{z_{k-1}}}\left[\frac{\Delta_{x_i}\Delta_{y_j}}{m_z^{(i,j,k-1)}}+\frac{\Delta_{x_{i-1}}\Delta_{y_j}}{m_z^{(i-1,j,k-1)}}+\frac{\Delta_{x_i}\Delta_{y_{j-1}}}{m_z^{(i,j-1,k-1)}}+\frac{\Delta_{x_{i-1}}\Delta_{y_{j-1}}}{m_z^{(i-1,j-1,k-1)}}\right]\Bigg\}
\end{aligned}
$$

where $\Delta_{\alpha_\beta} = \alpha_{\beta+1} - \alpha_\beta$, for $\alpha = x, y, z$ and the respective $\beta = i, j, k$, and $\Omega_{i,j,k} = \prod(\alpha_{\beta+1} - \alpha_{\beta-1})/2$ is the box volume concerning the node (i, j, k).

A more compact matrix notation can be used for the whole device Hamiltonian, representing the system of $N_x N_y N_z$ equations, with N_α the number of discretization nodes along the α direction:

$$
H_d = \left[\begin{array}{ccccc}
A_1 & B_{1,2} & 0 & 0 & \cdots \\
B_{2,1} & A_2 & B_{2,3} & 0 & \cdots \\
0 & B_{3,2} & A_3 & B_{3,4} & \cdots \\
\multicolumn{5}{c}{\dotfill} \\
\multicolumn{3}{c}{\dotfill} & B_{N_{x-1},N_x} & A_{N_x}
\end{array}\right] \tag{1.12}
$$

where each of the N_x block matrix A_i represents the discrete Hamiltonian for the i-th slice of the device along the transport direction including the additive potential energy, each diagonal block matrix $B_{i,i+1}$ describes the transfer integral between two points of adjacent slices, and $B_{i+1,i} = B^{\dagger}_{i,i+1}$. Each of the block matrices have dimension $(N_y N_z \times N_y N_z)$. The choice for the nearest neighbor approximation in the discretization of the Laplacian operator assures the block tridiagonal structure of global matrix H_d. The solution of the kinetic equations (1.3) in this representation is referred to as real-space (RS) approach. The RS approach is in general a heavy computational problem. Referring to the use of recursive-algorithm [8] for the calculation of both the retarded and the less-than Green's functions the most executed operations will be the inversion and the multiplication of matrices with the same leading dimension of the block matrices A and B. Depending on the lateral dimension of the device and on the chosen discretization step this operation can require large matrices storage and time-consuming executions. On the other hand, the RS approach allows for the exact inclusion of all the effects due to geometry variations (spatial fluctuations), correctly describing the inter-subband scattering for the entire set of 1-D subbands along the transport direction.

Concerning the calculation of the contacts self-energy functions, an analytical solution is not determined. A general expression can be derived from the truncation of the discrete Dyson's equations for a layered structure, and reads out [3]:

$$\Sigma^r_{\mathrm{lc}} = H_{d,\mathrm{lc}} g^r_{\mathrm{lc}} H^{\dagger}_{d,\mathrm{lc}} \ , \tag{1.13}$$

$$\Sigma^r_{\mathrm{rc}} = H_{d,\mathrm{rc}} g^r_{\mathrm{rc}} H^{\dagger}_{d,\mathrm{rc}} \ , \tag{1.14}$$

where $H_{d,\mathrm{lc}}(H_{d,\mathrm{rc}})$ is the Hamiltonian describing the coupling between the semi-infinite contact to the first(last) slice of the device, and $g^r_{\mathrm{lc}}(g^r_{\mathrm{rc}})$ is the retarded surface Green's function, describing the correlation between two points inside the lead when isolated from the device ($H_{d,\mathrm{lc}} = H_{d,\mathrm{rc}} = 0$). The surface Green's functions are expressed by a recursive relations and, independently of the basis set used, they can be found numerically through an iterative algorithm [9]. Alternative solutions, allowing for reduced computational costs, rely on proper basis set transformations [10], or, as for the case of the mode-space representation of the effective-mass problem, analytical formulations are possible.

In the presented formalism, the total electron density, resulting from the integration on a proper range of energies of the spectral density (1.5), reads out

$$n(x_i, y_j, z_k) = -\sigma_s \frac{i}{\Delta_x \Delta_y \Delta_z} \int \frac{\mathrm{d}E}{2\pi} G^{<}(i,i,j,j,k,k;E) \ , \tag{1.15}$$

where σ_s accounts for the spin degeneracy. The three different components of the current density results

$$\begin{aligned}
J_x(x_i, y_j, z_k) &= -\sigma_s \frac{q}{\hbar \Delta_y \Delta_z} \int \frac{\mathrm{d}E}{2\pi} \\
&\quad 2\Re\left[H_d(i, i+1, j, j, k, k) G^<(i+1, i, j, j, k, k; E)\right] \\
J_y(x_i, y_j, z_k) &= -\sigma_s \frac{q}{\hbar \Delta_x \Delta_z} \int \frac{\mathrm{d}E}{2\pi} \\
&\quad 2\Re\left[H_d(i, i, j, j+1, k, k) G^<(i, i, j+1, j, k, k; E)\right] \\
J_z(x_i, y_j, z_k) &= -\sigma_s \frac{q}{\hbar \Delta_x \Delta_y} \int \frac{\mathrm{d}E}{2\pi} \\
&\quad 2\Re\left[H_d(i, i, j, j, k, k+1) G^<(i, i, j, j, k+1, k; E)\right]
\end{aligned} \tag{1.16}$$

where $H_d(i, i', j, j', k, k')$ represents the discrete device Hamiltonian connecting the node (i, j, k) to (i', j', k'), and $\Re$ indicates the real-part of the resulting complex product. In the above equation the normalization of less-than Green's function for an uniform discretization step along each spatial direction is supposed, hence dropping out the point dependence of the discretization steps Δ_α for compactness.

1.3.2 Mode-Space approach

While still keeping all the relevant physics of the RS approach, the mode-space (MS) approach is introduced in order to reduce the computational costs for the solution of the transport problem. It consists in finding the real-space solution through its expansion on the basis set constituted by the wave functions solution of the 2-D Schrödinger problem defined at each slice of the device along the transport direction. As reported in [11], through simple arguments based on the application of the Kadanoff-Baym formalism [12], the real-space Green's function can be expressed as

$$G^<(i, i', j, j', k, k'; E) = \sum_{n,m} G^<_{nm}(i, i'; E) \phi^i_n(y_j, z_k) \phi^{i' \star}_m(y_{j'}, z_{k'}) \ , \tag{1.17}$$

where $G^<_{nm}(i, i'; E)$ is the mode-space counter part of the real-space Green's function, n and m sweep on all the transverse mode, and $\phi^i_n(y_j, z_k)$ is the n-th eigenfunction of the i-th slice of the device, verifying the problem

$$A_i \phi^i_n = (\varepsilon^i_n - 2t_x) \phi^i_n \ , \tag{1.18}$$

where A_i is the block matrix defined by (1.12), $2t_x$ is the on-site potential which, in the hypothesis of an uniform discretization and in absence of mass variations along the x axis, takes the usual form $t_x = \hbar^2/2m_x\Delta_x^2$, and ε^i_n is the n-th eigenvalue of the 2-D Schrödinger equation for the i-th slice. The succession of ε^i_n for $i = 1, 2, ..., N_x$ and for every n defines the 1-D subbands in the transport direction.

The sets of eigenfunctions, verifying at each slice the properties of orthonormality and completeness, can be grouped into the base change matrix

$$V = \begin{bmatrix} v_1 & 0 & 0 & 0 & \cdots \\ 0 & v_2 & 0 & 0 & \cdots \\ 0 & 0 & v_3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 0 & v_{N_x} \end{bmatrix} , \tag{1.19}$$

where $v_i = [(\phi_1^i(1), \ldots, \phi_1^i(N_y N_z))^{\mathrm{T}}, \ldots, (\phi_{N_y N_z}^i(1), \ldots, \phi_{N_y N_z}^i(N_y N_z))^{\mathrm{T}}]$ represents the complete set of eigenfunctions for each slice, and the matrix is supposed normalized $V^\dagger V = I$.

The advantages of the MS approach come from the use of a truncated expansion, retaining a subset of $N_m \ll N_y N_z$ modes, thus limiting the dimensions of the transport problem to be solved. The truncated basis set and the related matrix V_m is then used for the definition of the new mode-space Hamiltonian

$$H_{\mathrm{ms}} = V_m^\dagger H_d V_m \ , \tag{1.20}$$

that will conserve the block three-diagonal structure but with each block matrix having leading dimension $N_m \times N_m$. It can be easily found that each of the block matrices lying on the principal diagonal is diagonal, representing the on-site Hamiltonian for each of the N_m considered modes. The off-diagonal blocks, on the other side, are in general full matrices, deriving from the product ${v_i}^\dagger B_{ii\pm1} v_{i\pm1}$, and represent the coupling among the different transverse modes for every transition between two adjacent slices (correlation matrix). Using the previous assumptions, the discrete form of the device mode-space Hamiltonian operator associated to mode-space Green's function of the n-th mode on the i-th site is found to be:

$$-\, t_x \sum_m \left[c_i^{n,m} \delta_{i+1,j} + c_{i-1}^{n,m} \delta_{i-1,j} \right] + \left[2t_x + \varepsilon_i^n \right] \delta_{i,j} \ , \tag{1.21}$$

where $\delta_{i,j}$ is the Kronecker delta, and

$$c_i^{n,m} = \int \mathrm{d}y \mathrm{d}z \, \phi_i^n(y,z) \, {\phi_{i+1}^m}^*(y,z) \tag{1.22}$$

is the correlation between the n-th and the m-th mode for the i-th and $(i+1)$-th slices.

If the solution of the problem is carried out retaining all the informations connected to the correlation matrix, the method is usually referred to as Coupled Mode Space (CMS) approach, due to the fact that the solution is correctly describing the coupling effect between the considered transverse modes arising from the potential variations. An additional approximation is possible, discarding the correlation terms and forcing the correlation matrix to a diagonal matrix describing the propagation slice-by-slice of a single mode with no interactions with the others. This approach is referred to as Uncoupled Mode Space (UMS), and has the computational advantage of diagonalizing the set of equations to be solved in N_m uncoupled problems.

The transformation of the kinetic equations (1.3) is thus straightforward, keeping the same structure but with an obvious redefinitions of symbols moving from the real-space representation to its mode-space counterpart. Concerning the contact self-energy an analytical formulation can be found in the mode-space representation for the chosen contact model. It is worth noting that due the infinite repetition of the same confining potential an unique basis set is present in the contact and the correlations terms (1.22) are null, making no difference between the UMS and CMS description. Using a scalar notation for the entries of diagonal matrices, the left and right contact surface Green's functions for a propagating mode read out [13]

$$g^r = \begin{cases} (t_x)^{-1}(\chi + \sqrt{\chi^2 - 1}) & \text{when } |W_1| \leq 1 \\ (t_x)^{-1}(\chi - \sqrt{\chi^2 - 1}) & \text{when } |W_2| \leq 1 \ , \end{cases} \tag{1.23}$$

where $W_1 = -\chi + \sqrt{\chi^2 - 1}$ and $W_2 = -\chi - \sqrt{\chi^2 - 1}$, and

$$\chi = \begin{cases} \left[E - \Sigma^r_{\mathrm{ph}}(1; E) - 2t_x - \varepsilon_n(1)\right] /2t_x & \text{for lc} \\ \left[E - \Sigma^r_{\mathrm{ph}}(N_x; E) - 2t_x - \varepsilon_n(N_x)\right] /2t_x & \text{for rc} \ . \end{cases} \tag{1.24}$$

The effect of the electron-phonon interaction is, accordingly to the contact model, periodically repeated throughout the contacts. In the ballistic case, the contact self-energy for a single mode in a one-dimensional channel can be recast into a more intuitive expression. Inserting in (1.24) the discrete formulation for the 1-D effective-mass dispersion relation

$$E(k_x) = \varepsilon + 2t_x \left(1 - \cos\left(k_x \Delta_x\right)\right) \ , \tag{1.25}$$

and expressing the solution as a function of the wave vector k_x the resulting self-energy reads out

$$\Sigma^r_{\mathrm{lc/rc}} = -t_x e^{\mp i k_x^{\mathrm{lc/rc}} \Delta_x} \ , \tag{1.26}$$

where the phase factor recalls a plane wave propagating in the channel.

Finally, the electron and current densities are directly derived through the substitution of (1.17) in the respective real-space formulations:

$$n(x_i, y_j, z_k) = -\sigma_s \frac{i}{\Delta_x \Delta_y \Delta_z} \sum_{n,m} \int \frac{\mathrm{d}E}{2\pi} G^<_{nm}(i, i; E) \phi^i_n(y_j, z_k) {\phi^i_m}^\star(y_j, z_k) \ , \tag{1.27}$$

and from the first relation in (1.17)

$$\begin{aligned} J_x(x_i, y_j, z_k) \quad = \quad & -\sigma_s \frac{q}{\hbar \Delta_y \Delta_z} \sum_{n,m} \int \frac{\mathrm{d}E}{2\pi} 2\Re \left[H_d(i, i+1, j, j, k, k) \right. \\ & \left. \times \, G^<_{nm}(i+1, i; E) \phi^{i+1}_n(y_j, z_k) {\phi^i_m}^\star(y_j, z_k) \right] \end{aligned} \tag{1.28}$$

for the x component of the current density. The reduction to the UMS case is straightforward discarding all the off-diagonal terms of the mode-space Green's function. A relation analogous to (1.27) is found for the hole density, depending, in this case, on the greater-than Green's function, as discussed in further details in Sec. 2.3.

The integration on the confining plane of (1.28) gives the current along the transport direction, and in the case of purely ballistic transport, the Landauer-Büttiker formula is obtained [3]:

$$I = \sigma_s \frac{q}{\hbar} \int \frac{\mathrm{d}E}{2\pi} \mathrm{T}(E)(f(E-\mu_s) - f(E-\mu_d)) \tag{1.29}$$

where $\mathrm{T}(E) = \mathrm{Tr}\left[G^r \Gamma_{\mathrm{lc}} G^{r\dagger} \Gamma_{\mathrm{rc}}\right]$ is the transmission probability at the energy E, and Tr is the trace operator. In the limit of a linear transport regime (low drain-source biasing, V_{DS}) the linear approximation of (1.29) $I = GV_{\mathrm{DS}}$ holds, defining the conductance

$$G = \sigma_s \frac{q^2}{h} \int \mathrm{d}E \ \mathrm{T}(E) \left[-\frac{\partial f}{\partial E}(E-\mu_s) \right] \ . \tag{1.30}$$

Finally, some conclusions about the use of the MS approach are addressed. Several papers have been presented in literature treating the application of the mode-space approach and its comparison to the real-space result for 2-D structures like DG-MOSFETs [10, 11, 14]. The capability of the CMS approach to correctly reproduce the RS description of both current and charge density distributions was proved also in presence of potential constrictions (e.g. the presence of flared-out contacts in [11] or squeezed channels in [14]). The extension of the method to 3-D structures has been successfully demonstrated for the analysis of transport in SiNW-FETs in presence of potential deformations or fluctuations and also in the case of dissipative transport [15, 16, 17]. Applications are reported also in the case of CNT-FETs [18], where great advantage was demonstrated using the UMS approach on atomistic basis set.

Clear are the advantages in terms of reduced computational costs. If in the real-space approach the dependence on the number of discretization points can be evaluated as $O(N_x \times (N_y N_z)^3)$, where the dependence on $(N_y N_z)^3$ arises from the inversion of off-diagonal block matrices [8], it reduces to $O(N_x \times (N_m)^3)$ in the case of the mode-space approach. The total number of modes to be accounted for depends on the analyzed geometry. If no potential fluctuations are present and the UMS approximation holds, only few transverse modes can be considered depending on the confinement condition. On the other side, if important correlations are present, as in the case of rough channel/gate-oxide interfaces, a much larger number of modes (tens of modes) as to be accounted for, independently of their occupation, in order to have a good convergence to the real-space solution.

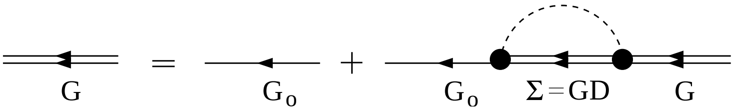

Figure 1.4: *Feymann diagram representation of the Dyson's equation of the electron-phonon interaction within the self-consistent Born approximation.*

1.4 The electron-phonon interaction

The electron-phonon interaction is included within a perturbative model within the self-consistent Born approximation. The calculation of the phonon self-energy can be derived by a diagrammatic representation of the Dyson's equation describing the second order contribution of the phonon interaction to the less-than Green's function [4], as depicted by the diagram of Fig. 1.4. Here, G represents the final perturbed Green's function (less-than Green's function in the following), being G_0 the unperturbed one. The interaction between the electron and phonon system is represented by a proper Σ function, confined between a starting and a final state (the circular vertex of the diagram).

Accordingly to the model, the phonon system is considered unperturbed by the interactions with the electron gas, hence the self-energy reported in Fig. 1.4 can be translated into

$$\Sigma_{\text{ph}}^{<,>} = G^{<,>} D^{<,>} \ , \tag{1.31}$$

where a real-space representation is supposed and the energy dependence has been omitted. $D^{<,>}$ represent the less-than and greater-than Green's functions of the unperturbed phonon bath, respectively. Their analytical formulations are calculated from the free-phonons Hamiltonian in the interaction picture, that in the harmonic approximation for the lattice vibrations reads out

$$\varphi_I(\mathbf{r},t) = \frac{1}{\sqrt{V}} \sum_q M_{\mathbf{q}} \left(a_{\mathbf{q}} e^{-i\omega_q t} + a_{\text{-}\mathbf{q}}{}^{\dagger} e^{i\omega_q t} \right) e^{i\mathbf{q}\cdot\mathbf{r}} \ , \tag{1.32}$$

where V is the volume of the sample, $\mathbf{q}$ is the wave vector of the phonons, ω_q is the corresponding angular frequency, $M_{\mathbf{q}}$ is the electron-phonon matrix element, and $a_{\mathbf{q}}$ and $a_{\mathbf{q}}^{\dagger}$ are the phonon creation and annihilation operators. Fourier-transforming the time-dependent operators, the self-energy functions at a given energy E are thus given by

$$\begin{aligned} \Sigma_{\text{ph}}^{<}(\mathbf{r}_1,\mathbf{r}_2;E) &= \int \frac{d\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)} |M_{\mathbf{q}}|^2 \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) \\ &\quad \times G^{<}(\mathbf{r}_1,\mathbf{r}_2;E \pm \hbar\omega_q) \end{aligned} \tag{1.33}$$

$$\begin{aligned} \Sigma_{\text{ph}}^{>}(\mathbf{r}_1,\mathbf{r}_2;E) &= \int \frac{d\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)} |M_{\mathbf{q}}|^2 \left(N_q + \frac{1}{2} \pm \frac{1}{2} \right) \\ &\quad \times G^{>}(\mathbf{r}_1,\mathbf{r}_2;E \mp \hbar\omega_q) \ , \end{aligned} \tag{1.34}$$

where N_q is the average phonon density at the energy $\hbar\omega_q$

$$N_q = \frac{1}{e^{(\hbar\omega_q/\kappa_B T)} - 1} \ . \tag{1.35}$$

The solution of the kinetic equations requires also the knowledge of the retarded self-energy, that can be calculated from the general relation [12]

$$\Sigma^r(E) = P \int \frac{\mathrm{d}\epsilon}{2\pi} \frac{\Gamma(\epsilon)}{E - \epsilon} - i\frac{\Gamma(E)}{2} \ , \tag{1.36}$$

where P is a principal value integral on the complex plane, and Γ, similarly to relations (1.10), is defined as

$$\Gamma(E) = i\left[\Sigma^{>}(E) - \Sigma^{<}(E)\right] \ , \tag{1.37}$$

for every point inside the device.

The real part of Σ^r, represented by the first term of the right side of (1.36), is a non-hermitian energy contribution giving a shift of the particle energy levels. The second term is associated to the the scattering rate due to the electron-phonon interaction, as stated in the previous section. The evaluation of (1.36) can be done analytically only for the case of elastic scattering while requiring a numerical evaluation of the principal value integral in the case of inelastic scattering process. This calculation can be computationally expensive due the necessity of the simultaneous knowledge of the Green's functions for every energy. Its contribution is in general omitted. The impact of this assumption has been investigated in [17], while, in [20] possible approximations for its inclusion in RTD are presented. In the following only the information connected to the imaginary part of the retarded self-energy will be retained within the approximation

$$\Sigma^r(E) \approx -i\frac{\Gamma(E)}{2} \ . \tag{1.38}$$

It is worth highlighting that through (1.36) or, equivalently, (1.38), a dependence of Σ^r on $G^<$ is introduced. This implies, that in presence of the electron-phonon interaction the first and the second equation in (1.3) are coupled through a non linear relation. As previously stated, in the ballistic case the two equations separately describe the dynamics and the statistical properties of the system. When the phonon scattering is included, hence, a self-consistent iterative solution of the two equations (1.3) with the phonon self-energy functions is required, in addition to the self-consistent loop with the Poisson equation. It is important to note that in the self-consistent Born approximation the conservation of the current flux is guaranteed along the entire device, verifying the fundamental fulfillment of the continuity equation.

The self-consistent Born approximation is surely the most widely used approach to include the electron-phonon interaction in the NEGF formalism. Approximations to the presented method or alternative approaches are in any case possible. Comparative studies between different approaches are present in literature. In [19] the common approximation of neglecting the coupling between the scattering states and their occupancy is discussed. It is shown how, although a clear computational advantage is obtained by artificially decoupling the kinetic equations, an unphysical result, neglecting the Pauli exclusion principle, is obtained. In [20] the multiple sequential scattering approach is presented as an alternative to the self-consistent Born approximation. The numerical advantages and physical limitations derived from the use of this model are discussed in relation to the analysis of a resonant-tunneling device.

In the following subsections the discrete form of (1.33) and (1.34) will be presented in the inelastic and elastic case for both silicon nanowires and carbon nanotubes, deriving the mode-space formulation from the real-space representation.

1.4.1 Phonon scattering in silicon nanowires

Transport of electrons in the conduction band of silicon is considered within the effective-mass approximation. There exist six minima in the conduction band, which can be grouped into three doubly degenerate valleys. Each valley will be labeled with an index $\nu = 1, 2, 3$. The case of both acoustic intra-valley and optical inter-valley scattering mechanisms will be presented.

In the case of acoustic intra-valley, phonon scattering is treated in the elastic approximation, considering $E \pm \hbar\omega_q \approx E$. Within the scalar assumption for the deformation potential Ξ the electron-phonon matrix element reads out [21]

$$|M_{\mathbf{q}}|^2 = \frac{\hbar \Xi^2 q}{2\rho u_l} \tag{1.39}$$

where $q = |\mathbf{q}|$, ρ is the silicon density, u_l is the sound velocity and a linear dispersion relation is assumed for the longitudinal acoustic mode. Due to the elastic approximation $N_q \approx N_q + 1 \approx k_B T/(\hbar u_l q)$ and the less-than Green's functions involved in the integral become independent of the phonon wave vector. Consequently, the real-space self-energy functions for each valley ν are expressed as

$$\Sigma^{<,>}_{\mathrm{ph},\nu}(\mathbf{r}_1, \mathbf{r}_2; E) = \frac{k_B T \Xi^2}{2\rho u_l^2} G^{<,>}_{\nu}(\mathbf{r}_1, \mathbf{r}_2; E)\delta(\mathbf{r}_1 - \mathbf{r}_2) \ , \tag{1.40}$$

where the Dirac delta is derived from the remaining integrand

$$\delta(\mathbf{r}_1 - \mathbf{r}_2) = \int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \ . \tag{1.41}$$

In the case of inter-valley transitions both acoustic and optical phonons can be involved. Six different type of transitions can be identified divided into g-type and f-type phonons, the former inducing transitions between the two degenerate valleys of same index ν, the latter transitions between valleys with different indexes. Approximating the momentum transfer for the j-th transition type as the momentum difference between the minimum energy states of the two valleys, $\mathbf{K}_j$, with a fixed energy transfer $\hbar\omega_j$, a formulation of the deformation potential independent of q is obtained

$$|M_{\mathbf{q}}|^2 = \frac{\hbar(D_t\mathbf{K}_j)^2}{2\rho\omega_j} \ . \tag{1.42}$$

Inserting (1.42) in (1.33) or (1.34), the j-th phonon self-energy for the ν-th valley is obtained

$$\begin{aligned} \Sigma^{<}_{\mathrm{ph},j,\nu}(\mathbf{r}_1,\mathbf{r}_2;E) &= \frac{\hbar(D_t\mathbf{K}_j)^2}{2\rho\omega_j}\left(N_j+\frac{1}{2}\pm\frac{1}{2}\right) \\ &\quad \times\sum_{\nu'=1}^{3} g_j^{\nu\nu'} G^{<}_{\nu'}(\mathbf{r}_1,\mathbf{r}_2;E\pm\hbar\omega_j)\delta(\mathbf{r}_1-\mathbf{r}_2) \end{aligned} \tag{1.43}$$

$$\begin{aligned} \Sigma^{>}_{\mathrm{ph},j,\nu}(\mathbf{r}_1,\mathbf{r}_2;E) &= \frac{\hbar(D_t\mathbf{K}_j)^2}{2\rho\omega_j}\left(N_j+\frac{1}{2}\pm\frac{1}{2}\right) \\ &\quad \times\sum_{\nu'=1}^{3} g_j^{\nu\nu'} G^{>}_{\nu'}(\mathbf{r}_1,\mathbf{r}_2;E\mp\hbar\omega_j)\delta(\mathbf{r}_1-\mathbf{r}_2) \end{aligned} \tag{1.44}$$

where $g_j^{\nu\nu'}$ represents the final subband degeneracy according to the selection rules

$$g_j^{\nu\nu'} = \begin{cases} \delta_{\nu\nu'} & \text{for } g\text{-type} \\ 2\,(1-\delta_{\nu\nu'}) & \text{for } f\text{-type} \ . \end{cases} \tag{1.45}$$

The presented scalar approximation for the deformation potentials is widely used and is found to give results for mobility within 3% agreement with more detailed Monte Carlo simulations [22].

Due to its repercussions on the simulation procedures and computational costs, it is interesting to note that the inclusion of optical modes causes a coupling between self-energies, and hence Green's functions, at energies differing of integer multiple of $\hbar\omega_j$. This implies the need of iteratively solving (1.3) simultaneously for a set of multiple energies.

The above relations have to be transformed from real-space to mode-space representation. The same procedure presented in the previous section with the projection over the basis set of slice eigenfunctions is used. Dropping out the superscripts and subscripts of the less-than and greater-than phonon self-energy, the n,m entry for the ν-th valley reads out

$$\begin{aligned} \Sigma_\nu^{n,m}(i,j;E) &= \int\!\!\int \mathrm{d}y\mathrm{d}z\mathrm{d}y'\mathrm{d}z' \\ &\quad \phi^{n\,\star}_{i,\nu}(y',z')\Sigma_\nu(x_i,x_{i'},y,y',z,z';E)\phi^{m}_{i',\nu}(y,z) \ . \end{aligned} \tag{1.46}$$

For the case of intra-valley acoustic phonons, inserting (1.40) in the above equation and making use of the expansion (1.17)

$$\Sigma_{\nu}^{n,n}(i,i;E) = \frac{k_B T \Xi^2}{2\rho u_l^2} \sum_m G_{\nu}^{m,m}(i,i;E) I_{m,\nu}^{n,\nu}(i,i) \ , \tag{1.47}$$

is obtained. All the contributions arising from the cross terms ($\Sigma^{n,m}$ for $n \neq m$) are neglected, reducing the self-energy formulation a diagonal matrix. This assumption has been validated numerically, noting that the overlap integral of multiple eigenfunctions associated to the off-diagonal terms of the Green's function in the evaluation of (1.47) give a negligible contribution compared with the terms accounted for. Hence, $I_{m,\nu}^{n,\nu'}$ represents the usual form factor

$$I_{m,\nu}^{n,\nu'}(i,i) = \int \mathrm{d}y\mathrm{d}z \left|\phi_{i,\nu'}^{n}(y,z)\right|^2 \left|\phi_{i,\nu}^{m}(y,z)\right|^2 \ . \tag{1.48}$$

The same formulation holds for both the less-than and greater-than function with a proper redefinition of symbols.

In the case of inter-valley scattering a similar result is obtained:

$$\begin{aligned} \Sigma_{j,\nu}^{n,n}(i,i;E) \quad &= \quad \frac{\hbar (D_t \mathbf{K}_j)^2}{2\rho\omega_j} \left(N_j + \frac{1}{2} \pm \frac{1}{2} \right) \\ & \quad \times \sum_{m,\nu'} g_j^{\nu\nu'} G_{\nu'}^{m,m}(i,i;E \pm \hbar\omega_j) I_{m,\nu}^{n,\nu}(i,i) \ , \end{aligned} \tag{1.49}$$

for each of the six type of scattering and and where the less-than and greater-than symbols have been dropped out.

It is important to recall that in the presented derivation of the phonon scattering self-energy functions a bulk model has been assumed for the phonon system. Effects of quantization on the phonon dispersion relation due to the confinement on the transverse plane are neglected. Simulation studies on the effect of confinement for phonon scattering in silicon nanowires have been recently presented in [22, 23]. In [23] the role of acoustic confined phonons has been investigated, additionally to surface-roughness scattering, through Monte Carlo simulations, showing a maximum 10% mobility reduction with respect to a bulk model. A TB model for electron and hole energy dispersion relation in small nanowires (diameter less than 3 nm) is used in scattering rates and mobility calculations in [22]. Intrasubband confined acoustic phonon and bulk optical phonon scattering are considered. As the wire diameter increases the difference between mobility due to confined and bulk acoustic phonons decreases, and bulk phonons is found to be a relatively good approximation for some of the larger diameter freestanding SiNWs. It is noted that, although the investigation of confinement effect in SiNW electron-phonon interaction is a fundamental topic requiring further investigations, the bulk approximation is still widely used in both semi-classical [24] and quantum models [16, 17], giving a valid description of all qualitative aspects of this scattering mechanisms. Additional details on the simulated scattering mechanisms and the specific values for scattering parameters will be given in Chapter 3.

1.4.2 Phonon scattering in carbon nanotubes

Carbon nanotubes are ideal 1-D systems where the charge is localized on the tube surface. Transport occurs along the tube longitudinal direction on 1-D subbands deriving from the quantization on the circumferential direction of the graphene energy dispersion relation, as presented in Sec. 2.1. The general real-space formulation of the problem is the same as presented for the silicon case.

The electron-phonon interaction is governed by energy and momentum conservation. In the case of CNTs the angular quantization affecting both electron and phonon energy dispersion can not be neglected. Electrons scattered within the same subband have no change in their angular momentum. Thus, such a mechanism is assisted by a zone-center phonon with null angular momentum, $l = 0$ [25]. In the case of subbands with different angular momentum m and m', the scattering has to be mediated by a zone-boundary phonon with angular momentum $l = |m - m'|$ [25], as can be derived from simple momentum conservation rules based on the zone-folding method [26]. An analysis of the selection rules and scattering mechanisms for different CNTs is reported in [27]. Considering the fulfillment of the required momentum conservation, the uncoupled mode space representation of the scattering self-energy will be derived for the m-th subband. It is worth noting that, in the case of CNT, the analysis proposed in Chapter 2 will be restricted to the case of devices with a Gate-All-Around architecture and in absence of defects or deformation of the tube. Hence, the UMS approach will be used, being capable of perfectly reproduce the real-space solution in terms of few transverse modes [18].

For the case of elastic scattering, as approximation of the acoustic phonon assisted transitions, a linear energy dispersion is assumed with appreciable occupation only for phonons with momentum close to zero. Starting from a relation similar to (1.40), the mode-space formulation of the less-than and greater-than self-energy for the m-th subband reads out

$$\Sigma_m^{<,>}(i,i;E) = \frac{k_B T \tilde{K}_a^2}{\rho_{1D} v_a^2} G_{m'}^{<,>}(i,i;E) \ , \tag{1.50}$$

where $\tilde{K}_a^2$ is the electron-phonon matrix element, ρ_{1D} the one-dimensional mass-density, v_a the sound velocity, and m' is the index of the subband verifying the selection rules for the transverse momentum. Note that, in the notation, the second subband index has been removed since, in UMS approach, the problem is reduced to a set of independent equations. In addition, the form factor function is identically unitary for every couple (m, m') since, due to the Gate-All-Around architecture, the CNT eigenfunctions are circular harmonics $\exp(im\varphi)$, with φ the radial coordinate.

For the interaction with inelastic optical phonons, a isotropic scattering mechanism is supposed independent of the phonon wave vector q, leading to relations analogous to (1.43) and (1.44). The mode-space formulations are thus given by:

$$\begin{aligned}\Sigma_m^{<}(i,i;E) &= \frac{\hbar|K_0|^2}{2\rho_{1D}\omega_0}\left(N_0+\frac{1}{2}\pm\frac{1}{2}\right)\\ &\quad G_{m'}^{<}(i,i;E\pm\hbar\omega_0)\end{aligned} \qquad (1.51)$$

$$\begin{aligned}\Sigma_m^{>}(i,i;E) &= \frac{\hbar|K_0|^2}{2\rho_{1D}\omega_0}\left(N_0+\frac{1}{2}\pm\frac{1}{2}\right)\\ &\quad G_{m'}^{>}(i,i;E\mp\hbar\omega_0)\end{aligned} \qquad (1.52)$$

where $\hbar\omega_0$ is the phonon energy with matrix element $|K_0|^2$. It is worth noting the obtained self-energy functions, in both the silicon and CNT case, are expressed in a local formulation, with a dependence of the phonon self-energy at the node (i,j,k) on only the Green's function at the same node. Observations on the meaning and the implication of this constraint are presented in [25].

1.5 The Poisson equation

In the Hartree approximation for the electron-electron interaction, the electrostatic problem is completely solved by the Poisson equation, reported in (1.2) for a 3-D domain with non uniform medium (position dependent dielectric constant). Due to the coupling with the Schrödinger equation through the carrier density, the total charge is non linearly dependent on the potential itself and an explicit formulation of ρ as a function of ϕ is not possible. The system of $N_xN_yN_z$ equations representing the discretization of the Poisson equation can be thus generally expressed in the form:

$$\mathbf{F}\left[\boldsymbol{\phi},\boldsymbol{\rho}(\boldsymbol{\phi})\right]=0\ , \qquad (1.53)$$

where $\boldsymbol{\phi}=[\phi_1,\ldots,\phi_{N_xN_yN_z}]$ and $\boldsymbol{\rho}=[\rho_1,\ldots,\rho_{N_xN_yN_z}]$ are the vectors representing the discretized potential and charge, respectively. In order to have better convergence properties for the self-consistent problem, a Newton-Raphson scheme is adopted, solving (1.53) in terms of the correction vector $\delta\boldsymbol{\phi}$ of a trial solution $\boldsymbol{\phi}^0$

$$\mathbf{F}\left[\boldsymbol{\phi}^0+\delta\boldsymbol{\phi}\right]=\mathbf{F}\left[\boldsymbol{\phi}^0\right]+\mathbf{J}\delta\boldsymbol{\phi}=0\,, \qquad (1.54)$$

where $\mathbf{J}$ represents the Jacobian matrix of the partial derivatives of $\mathbf{F}$ with respect to the components of $\boldsymbol{\phi}$.

The definition of the Jacobian matrix $\mathbf{J}$ is a non trivial part of the problem due to the complex dependence of the electron and hole density on the electrostatic potential in a quantum formulation of the problem. First of all, a local dependence of the charge on the potential is assumed, expressing the response of the charge to the potential correction through a point-by-point first-order expansion

$$\rho_i(\phi_i) = \rho_i(\phi_i^0) + \left.\frac{\partial \rho}{\partial \phi}\right|_i \delta\phi_i \ , \ \text{for } i = 1, \ldots, N_x N_y N_z \ . \tag{1.55}$$

Several solutions have been proposed to calculate the set of partial derivatives $\partial\rho/\partial\phi$, ensuring different convergence properties. One possibility is that of introducing an additional transformation for the unknown of the problem in order to have a functional dependence of the charge on the potential giving both a physic based relation and an explicit formulation of the partial derivatives. In the analysis of SiNW-FETs, the adopted choice has been that of expressing the relation between electron density and potential exploiting the semi-classical relation:

$$n_i(\phi_i) = N_c \, \mathcal{F}_{\frac{1}{2}} \left(q \, \frac{\phi_i - \Phi_{n_i}}{k_B T} \right) \ , \ \text{for } i = 1, \ldots, N_x N_y N_z \tag{1.56}$$

being, N_c the effective density of states, $\mathcal{F}_\eta$ the η-order Fermi-Dirac integral, and $\mathbf{\Phi}_n = [\Phi_{n_1}, \ldots, \Phi_{n_{N_x N_y N_z}}]$ an auxiliary variable, physically related, at quasi-equilibrium conditions, to the electron quasi-Fermi potential and without a clear physical significance under off-equilibrium conditions [28]. For the solution of Poisson equation in the analysis of CNT-FETs, an analogous procedure is introduced but with the use of the relation for a 1-D system and doing a separate transformation for the electron and hole density.

Through the transformation (1.56) a numerical evaluation of the Jacobian is possible at every iteration step. The introduction of the new auxiliary variable $\mathbf{\Phi}_n$, based on an arbitrary transformation, requires an additional iterative procedure where, at each step, the charge and its derivatives with respect to the potential are consistently updated while $\mathbf{\Phi}_n$ is kept fixed. The cost of this additional self-consistent loop is, any case, highly affordable due to the fast resolution of the linear system associated to the Poisson equation and to the good convergence of the additional iterative procedure. At each iteration step, the potential is updated with the current correction vector as:

$$\phi_i^k = \phi_i^{k-1} + \delta\phi_i^k \,, \tag{1.57}$$

where k is the iteration counter, and $i = 1, \ldots, N_x N_y N_z$.

From the numerical point of view, as for the Schrödinger equation, the box-integration method has been used for the discretization of the problem. Neumann boundary conditions are applied in the contact area and in the oxide regions not covered by the gate. This means that the solution in the contacts is a uniform potential, whose position with respect to the Fermi level of the reservoir is fixed by the charge neutrality constraint. Dirichlet boundary conditions are applied at the gate electrode. Finally, a gradient-conjugate procedure has been used for the solution of the linear system obtained by the discretization, with a packed storing of the system matrix, in order to exploit its sparsity proprieties to have a faster solution and a less occupation of memory. In Chapter 3 the treatment of the exchange-correlation correction to the electrostatic potential will be introduced in an approximated formulation based on the Density-Functional-theory.

1.6 Summary

In this Chapter a review of the NEGF formalism was given. The fundamental equations for the description of the transport properties of a MOSFET have been presented in the steady-state condition. The mode-space approach has been derived from the real-space representation within the coupled and uncoupled modes approximation. The inclusion of the electron-phonon interaction has been discussed with particular reference to its application to CNTs and SiNWs.

Bibliography

[1] R. Landauer, “Spatial variation of Currents and Fields due to Localized Scatterers in Metallic Conduction”, *IBM J. Res. Dev.*, vol. 1, p. 233, 1957. [cited at p. 5]

[2] M. Büttiker, “Four-Terminal Phase-Coherent Conductance”, *Phys. Rev. Lett.*, vol. 57, p. 1761, 1986. [cited at p. 5]

[3] S. Datta, *Electronic transport in Mesoscopic Systems*, Cambridge University Press, Cambridge, 5th ed., 2003. [cited at p. 5, 6, 7, 10, 14]

[4] A.L. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems*, Dover, New York, 2003. [cited at p. 6, 7, 15]

[5] G.D. Mahan, *Many-particle physics*, Kluwer, New York, 3rd ed., 2000. [cited at p. 6]

[6] Mathieu Luisier, *Quantum Transport Beyond the Effective Mass Approximation*, Ph.D. Thesis, Swiss Federal Institute of Technology, Zürich, 2007. [cited at p. 7]

[7] R. Lake and S. Datta, “Nonequilibrium Green’s-function method applied to double-barrier resonant-tunneling diodes”, *Phys. Rev. B*, vol. 45, no. 12, pp. 6670–6685, 1992. [cited at p. 8]

[8] G. Klimeck, *“Numerical Aspects of NEGF: The Recursive Green Function Algorithm”*, http://www.nanohub.org/resources/165/, 2005. [cited at p. 10, 14]

[9] M. P. Lopez Sancho, J. M. Lopez Sancho, and J. Rubio, “Higly convergent schemes for the calculation of bulk and surface Green functions”, *J. Phys. F: Met. Phys.*, vol. 15, p. 851, 1985. [cited at p. 10]

[10] R. Venugopal, Z. Ren, S. Datta, and Mark Lundstrom, “Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches”, *J. of Appl. Phys.*, Oct. 1, 2002, Vol. 92, No. 7, pp. 3730-3739. [cited at p. 10, 14]

[11] M. Luisier, A. Schenk, and W. Fichtner, “Quantum transport in two- and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium Green’s function formalism”, *Jour. Appl. Phys.*, vol. 100, p. 043713,2006 [cited at p. 11, 14]

[12] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*, Benjamin, New York, 1962. [cited at p. 11, 16]

[13] M. J. McLennan, Y. Lee, and S. Datta, “Voltage Drop in Mesoscopic Systems: A Numerical Study using a Quantum Kinetic Equation,” *Phys. Rev. B*, vol. 43, pp. 13846-13884, 1991. [cited at p. 13]

[14] R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, “Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors”, *J. Appl. Phys.*, pp. 292-305, 2004. [cited at p. 14]

[15] J. Wang, E. Polizzi, A. Ghosh, S. Datta, M. Lundstrom, “ Theoretical investigation of surface roughness scattering in silicon nanowire transistors”, *Apll. Phys. Lett.*, vol. 87, p. 043101, 2005. [cited at p. 14]

[16] S. Jin, Y. J. Park, and H. S. Min, “A three-dimensional simulation of quantum transport in silicon nanowire transistor in presence of electron-phonons interactions” *J. Appl. Phys.*, vol. 99, p. 123719, 2006. [cited at p. 14, 19]

[17] M. Frey, A. Esposito, and A. Schenk “Simulation of intravalley acoustic phonon scattering in silicon nanowires”, *European Solid state device Research Conference ESSDERC 2008*, pp. 258-261, 2008 [cited at p. 14, 16, 19]

[18] J. Guo, S. Datta, M. Lundstrom, M. P. Anantram, “Toward multiscale modeling of carbon nanotube transistors”, *Int. J. Multiscale Comput. Eng.*, vol. 2, p. 60, 2004. [cited at p. 14, 20]

[19] T. Kubis and P. Vogl, “Self-consistent quantum transport theory: Applications and assessment of approximate models” *J. Comput. Electron.*, vol. 6, p. 183, 2007. [cited at p. 17]

[20] R. Lake, G. klimeck, R. C. Bowen, and D. Jovanovic, “Single and multiband modeling of quantum electron transport through layered semiconductor devices” *J. Appl. Phys.*, vol. 81, p. 7845, 1997. [cited at p. 16, 17]

[21] C. Jacoboni and L. Reggiani, “The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials”, *Reviews of Modern Physics*, vol. 55, no. 3, pp. 645-705, 1983. [cited at p. 17]

[22] A. K. Buin, A. Verma, and M. P. Anantram, “Carrier-phonon interaction in small cross-sectional silicon nanowires”. *J. Appl. Phys.*, vol. 104, p. 053716, 2008. [cited at p. 18, 19]

[23] E. B. Ramayya, D. Vasileska, S. M. Goodnick, and I. Knezevic, "Electron Transport in Silicon Nanowires: The Role of Acoustic Phonon Confinement and Surface Roughness Scattering", *J. Appl. Phys.*, vol. 104, p. 063111, 2008. [cited at p. 19]

[24] S. Jin, M.V. Fischetti, and T.-W. Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band parabolicity", *J. Appl. Phys.*, vol. 102, p. 083715, Oct. 2007. [cited at p. 19]

[25] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, "Nonequilibrium Green's Function Treatment of Phonon Scattering in Carbon-nanotube Transistor", *IEEE Trans. on Electr. Dev.*, vol. 54, p. 2339, September 2007. [cited at p. 20, 21]

[26] E. B. Barros *et al.*, "Review on the symmetry-related properties of carbon nanotubes", *Physics Reports*, vol. 431, pp. 261–302, 2006. [cited at p. 20]

[27] G. Pennington and N. Goldsman, "Semiclassical transport and phonon scattering in semiconducting carbon nanotubes", *Phys. Rev. B*, vol. 68, no. 4, p. 045426, 2003. [cited at p. 20]

[28] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, "nanoMOS 2.5: A Two-Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs" *IEEE Trans. on Electr. Dev.*, vol. 50, pp. 1914–1925, 2003. [cited at p. 22]

Part II

APPLICATION TO THE FETs ANALYSIS

Chapter 2

Transport analysis in CNT based FETs

Carbon-nanotube field-effect-transistors (CNT-FETs) have received considerable attention in the past few years because of their excellent performance. Experiments with quasi-planar gate geometries have demonstrated near-ballistic transport in p-type CNT-FETs and on-currents that exceed those of silicon-based transistors. An additional enhancement is expected from the Gate-All-Around geometry, because it provides a better electrostatics control, thus pushing forward the ultimate scaling limits.

Due to the novelty of some electronic properties of CNTs, numerical analyses are still of basic importance to understand the limiting factors in the transport performance of such new devices. As semiconducting CNTs have experimentally demonstrated typical band gaps smaller than 1 eV, quantum effects like band-to-band tunneling are rising as being of fundamental importance for a realistic description of CNT-FETs. Theoretical studies have focused also on the evaluation of the impact of the electron-phonon interaction, aiming at understanding if CNTs can really behave as ballistic semiconductors.

In this framework, full-quantum models have rapidly emerged as favored candidates for the transport analysis of CNT-FETs. Due to the particular nature of CNTs, tight-binding models have been demonstrated to accurately reproduce the electronic properties of nanotubes even when a single-parameter model is used. It is in any case of interest to search for less expensive computational models capable of maintaining a realistic description of transport, like, for instance, simple extensions of the effective-mass approximation.

In addition, computational studies are still required in order to evaluate the best architectural choices for CNT-FETs and alternatives to the standard

MOSFET-like architecture. Exploiting the small band-gap energy and the low effective masses in new device concepts like tunneling-FETs, can extend the usage of CNT-based devices from high-performance to low-power applications.

This chapter reports on the application of a full-quantum transport model in the analysis of CNT-based FETs and tunneling devices. In Sec. 2.1 a description of the fundamental electronic properties of carbon nanotubes is given. The two device architectures that will be extensively studied in terms of fundamental transport properties and optimization of the geometrical and technological features are presented in Sec. 2.2. An extension of the parabolic effective-mass model, which gives a realistic description of transport in CNT-FETs in both ballistic and dissipative conditions is presented in Sec. 2.3. A validation of the adopted approach is presented with direct reference to tight-binding models. The first device, a MOSFET-like FET, is analyzed in Sec. 2.4, focusing mainly of the effect of the electron-phonon interaction in the scaling process. The second one, a tunneling-FET, is finally presented in Sec. 2.5, where a computational study of the ultimate scaling possibility of this device is presented focusing on low-operating-power applications.

2.1 Electronic properties of carbon nanotubes

The nanotube can be conceptually viewed as a rolled-up graphene sheet. Thus the electronic properties of a CNT can be directly derived from the ones of the two-dimensional graphene. In particular, the aim of this section is to recall the main characteristics of the nanotube 1-D energy dispersion relation which can be derived as quantization of the 2-D $E(\mathbf{k})$ of the graphene sheet along the circumferential direction of the nanotube [1].

The 2-D graphene lattice in the real space is depicted in Fig. 2.1. It can be obtained as the infinite repetition of the the unit cell constituted by the atoms labeled A and B along the directions identified by the translation vector $\mathbf{l} = n\mathbf{a}_1 + m\mathbf{a}_2$, with n and m integers and $\mathbf{a}_1$ and $\mathbf{a}_2$ the lattice characteristic vectors. The two characteristic vectors are expressed as

$$\begin{aligned} \mathbf{a}_1 &= a(\frac{\sqrt{3}}{2}, \frac{1}{2}) \\ \mathbf{a}_2 &= a(\frac{\sqrt{3}}{2}, -\frac{1}{2}) \ , \end{aligned} \tag{2.1}$$

where $a = \sqrt{3}a_{\rm cc}$ and $a_{\rm cc} = 1.42$ Å is the carbon-to-carbon bonding distance.

In a TB model, the single-electron Hamiltonian can be expressed as an expansion on a localized atomic orbitals set. Due to the sp^2 hybridization of the carbon molecular orbitals in a graphene lattice, and neglecting the spin-orbit coupling, the electronic properties of this material are well reproduced, in particular around the Fermi energy, accounting only for the interactions of the nearest

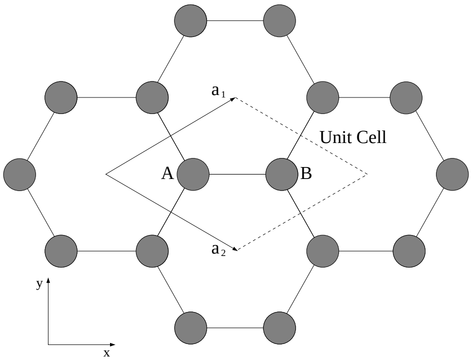

Figure 2.1: *Representation of the graphene unit cell. The two carbon atoms of a single cell are labeled as A and B. The lattice characteristic vectors are reported.*

p_z orbitals. Hence, the discrete component of the Hamiltonian describing the interaction between two atoms, which reads

$$H_{\alpha,\beta}(\mathbf{l}_n - \mathbf{l}_m) = \langle \alpha, n \left| H \right| \beta, m \rangle \ , \tag{2.2}$$

where α and β identify the atoms in the unit cell n and m, will be characterized by a single empirical parameter $v_{pp\pi}$, representing the bonding energy for two adjacent atoms.

Within the assumption of an orthogonal basis set, the generalized eigenvalue problem or secular equation for the n-th branch of the dispersion relation, $E_n(\mathbf{k})$, can be written as

$$\sum_\beta \left\{ \sum_m H_{\alpha,\beta}(\mathbf{l}_n - \mathbf{l}_m) e^{i\mathbf{k}\cdot(\mathbf{l}_n - \mathbf{l}_m)} \right\} c_\beta = E_n(\mathbf{k}) c_\alpha \ , \tag{2.3}$$

where c_α is the corresponding wave function.

In the case of a graphene sheet, in the nearest-neighbor approximation, the total Hamiltonian reads out

$$H = v_{pp\pi} \begin{pmatrix} 0 & \left(1 + 2e^{-i\frac{\sqrt{3}}{2}ak_x} \cos\left(\frac{a}{2}k_y\right)\right) \\ \left(1 + 2e^{i\frac{\sqrt{3}}{2}ak_x} \cos\left(\frac{a}{2}k_y\right)\right) & 0 \end{pmatrix} \ , \tag{2.4}$$

where a null value has been considered for the on-site energy of the single atom, giving just a stiff shift of the band structure, and where $\mathbf{k} = (k_x, k_y)$ is the 2-D wave vector.

In the case of the Hamiltonian (2.4) an analytical solution of (2.3) is found

$$E(k_x, k_y) = \pm v_{pp\pi} \sqrt{1 + 4\cos(\frac{a}{2}k_y)\cos(\frac{\sqrt{3}a}{2}k_x) + 4\cos^2(\frac{a}{2}k_y)} \ , \tag{2.5}$$

giving the two branches of the graphene energy dispersion relation corresponding to the π-energy bonding band and the $\pi^\star$-energy anti-bonding band. The corresponding energy surface is presented in Fig. 2.2 where (left) the 3-D plot and

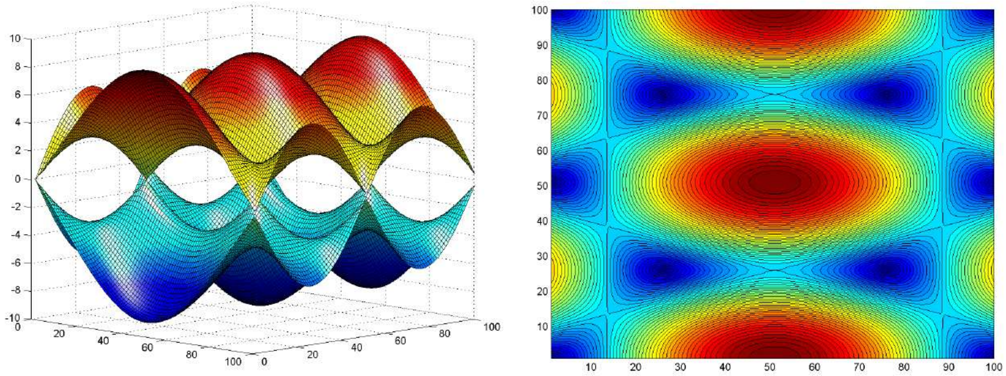

Figure 2.2: *(Left) Two-dimensional graphene dispersion relation and (right) its contour plot. The plots are given in arbitrary units. Only the π and $\pi^\star$ branches of the $E(\boldsymbol{k})$ are reported, calculated in the nearest-neighbor TB approximation neglecting the overlap between atomic orbitals. From the contour plot is evident the hexagonally shaped Brillouin zone of graphene.*

(right) the contour plot are reported. From the contour plot, the hexagonal shape of the graphene Brillouin zone is clearly evidenced. Within the approximation of an orthogonal set of atomic orbitals, a symmetric structure of the two branches is found. On the other side, the inclusion of an overlap matrix will result in a deformation of the bonding band [1]. Focusing only on the analysis of transport properties, the symmetric treatment of carriers in conduction and valence band has found to have just small quantitative effects on such quantities.

From the plot of the graphene 2-D $E(\mathbf{k})$ it can be seen that energy valleys with zero band gap are located at the corners of the Brillouin zone, referred to as Fermi points. It can be shown that a first order expansion of the energy dispersion relation around the Fermi points leads to a conical shape of the $E(\mathbf{k})$ with carriers acting similarly to massless relativistic particles [2]. Since the density of states at the Fermi level is zero, the 2-D graphite sheet is a zero band gap semiconductor.

The usual way to approach the calculation of the band structure of a carbon nanotube is to describe the CNT as rolled-up graphene sheet applying the zone-folding (ZF) technique to the tight-binding model. The periodic boundary condition only allows for quantized wave vectors along the circumferential direction, which generates 1-D bands [3]. The circumferential section of the nanotube is defined through a given translational vector, hereafter named chiral vector $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$. Beside, the translational vector $\mathbf{T}$, parallel to the axis of the tube and orthogonal to the chiral vector in the unrolled graphene lattice, is defined to be the unit vector of the 1D carbon nanotube. The rectangle enclosed by $\mathbf{C}_h$ and the vector $\mathbf{T}$ thus identifies the unit cell of the carbon nanotube, as reported in Fig. 2.3.

The imposition of the periodic boundary conditions implies that the reciprocal

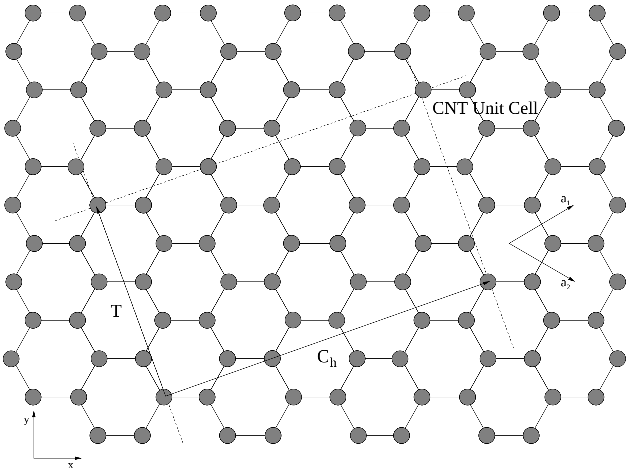

Figure 2.3: *Representation of the unrolled CNT unit cell. The chiral ($\boldsymbol{C}_h$) and translational ($\boldsymbol{T}$) vectors are reported.*

lattice vector $\mathbf{k}_1$ corresponding to $\mathbf{C}_h$ is quantized

$$\mathbf{k}_1 \cdot \mathbf{C}_h = 2\pi q \ , \tag{2.6}$$

where $q = 1, \ldots, N_{\text{cell}}$ is an integer, with N_{cell} the number of primitive graphene unit cells included in the CNT unit cell. The wave vector $\mathbf{k}_2$ reciprocal of $\mathbf{T}$ will remain continuous identifying the 1-D nanotube Brillouin zone (a line parallel to $\mathbf{k}_2$ itself). Hence, the energy dispersion relation for the CNT consists of N_{cell} couples of one-dimensional energy bands within the Brillouin zone of length $2\pi/|\mathrm{T}|$, where each band is a cut of the graphene dispersion relation on the lines $(\mathbf{k}\mathbf{k}_2/|\mathbf{k}_2| + i\mathbf{k}_1)$, each cut being folded in the first Brillouin zone of the nanotube.

The geometrical and the electronic properties of the nanotube are directly connected to the chosen orientation of the chiral and translational vectors, i.e. by the choice of the couple of integers (n, m), defining the tube chirality. It is important to note that proper combinations of n and m give cuts of the graphene energy dispersion relation which pass through the Fermi points. In these cases, the obtained 1-D subband will be characterized by a linear dispersion relation and a zero band gap. The obtained tube will show metallic properties with a 1-D density of state constant around the crossing point. By means of geometrical observations [1], it can be found that the (n, m) combinations verifying the condition

$$[n - m]_{\text{mod}_3} = 0 \tag{2.7}$$

select metallic nanotubes. In particular nanotubes with $n = m$, identified as *armchair* nanotubes, will be always metallic.

Connected to the interest of using nanotubes as channel material for FETs, CNTs showing a semiconducting behavior will be addressed in the following.

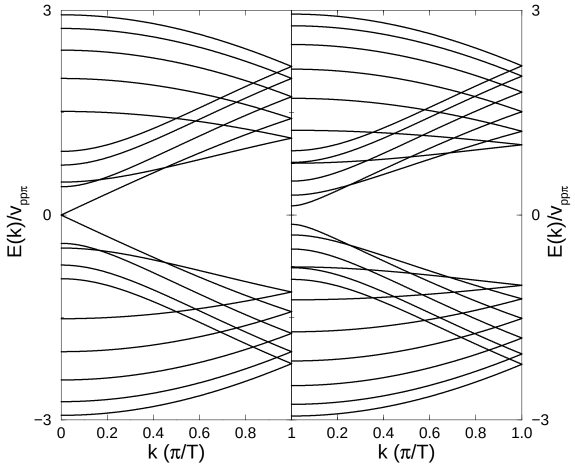

Figure 2.4: *One-dimensional energy dispersion relation for (left) a metallic* $(9,0)$ *and (right) a semiconducting* $(11,0)$ *zigzag CNT. The TB parameters are:* $v_{pp\pi} = 2.7$ *eV,* $a_{\rm cc} = 1.42$ *Å.*

In particular, the reference structure will be that of the family of achiral tubes defined as *zigzag*, identified by the chiral indexes $(n,0)$, with $[n]_{\rm mod_3} \neq 0$. It is also important to remind that the presented result is based on an ideal model which neglects any deformation effect induced by the graphene sheet carbon bondings as it is "transformed" in a closed tube. In a real case, variation of the bonding lengths can occur, in particular when the radius of the tube is reduced, due to a curvature effect. It has been reported in [4] the result of a Density-Functional-Theory (DFT) study for a vast class of carbon nanotubes, showing how, the semiconducting zigzag tubes become all metallic for $n < 11$. On the contrary, the smallest armchair tubes show the formation of a band gap. In any case, the ZF model applied to a TB description is found to give appreciable results for tube diameters larger than 1.0 nm.

Examples for the obtained 1-D energy dispersion relations are reported in Fig. 2.4 for a metallic and a semiconducting zigzag CNT (details given in caption). In the case of the $(n,0)$ zigzag CNT, the unitary cell contains $2n$ primitive graphene cells. Hence, $4n$ branches of the energy dispersion are present, with $2n$ valence bands and $2n$ conduction bands. Double degeneracy for the lowest conduction (higher valence) bands occurs. The analytical expression (2.5) for of the m-th subband reads out [1]:

$$E_m(k) = \pm v_{pp\pi}\sqrt{1 + 4\cos(\frac{\sqrt{3}ka}{2})\cos(\frac{m\pi}{n}) + 4\cos^2(\frac{m\pi}{n})}\ , \qquad (2.8)$$

where the first Brillouin zone extends in the interval $-\pi/\sqrt{3}a < k < +\pi/\sqrt{3}a$. The minimum of each subband is located at $k = 0$, hence the dependence on n

of the energy gap between each couple of conduction and valence band is

$$Eg_m = \sqrt{3} a v_{pp\pi} \frac{2\pi}{na} \left(m - \frac{2n}{3} \right) . \tag{2.9}$$

Equation (2.9) clearly shows that a metallic zigzag CNT has a zero band gap for the quantum number $m = 2n/3$. If n is not an integer multiple of 3, the minimum gap value is obtained for $m = \lceil 2n/3 \rceil$ and it reduces to

$$E_g = \frac{\sqrt{3} a v_{pp\pi}}{3} \frac{2\pi}{na} = \frac{2 a_{cc} v_{pp\pi}}{d_{\text{CNT}}} \approx \frac{0.8}{d_{\text{CNT}}} \text{eV} \cdot \text{nm} , \tag{2.10}$$

where $d_{\text{CNT}} = n\sqrt{3} a_{cc}/\pi$ is the tube diameter, expressed in nanometers.

The dependence of the electronic properties of CNTs on the chirality and, for a fixed chirality, on the diameter have been experimentally proved. Semiconducting single walled nanotubes have been reported with diameters in the range of $1.0-1.4$ nm with excellent properties in terms of band gap [5, 6, 7]. Although a clear experimental control on tube chirality has not been reported, a fitting of the experimental results is possible using the previous relation derived for zigzag CNTs in a range of band gap values interesting for electronic applications [8]. For this reason in the following, zigzag carbon nanotubes with a diameter $d_{\text{CNT}} \approx 1.0$ nm and $E_g \approx 0.8$ eV will be used as reference devices.

2.2 Device architecture

A 3-D sketch of the CNT-FET architecture used as device model is reported in Fig. 2.5. An ideal Gate-All-Around configuration is chosen with a cylindrical oxide layer clothing the tube for its entire extension. This architectures ensure the best electrostatic control achievable minimizing short channel effects and optimizing the sub-threshold slope.

Various experimental realizations have been reported for CNT-FET with different contact typologies. Even if Schottky barrier transistors (SBCNT-FETs) are presented as the simplest realization of functional FET devices [9, 10, 11], doped contacts have been reported and studied. Both electrostatic [12] and chemical doping [13, 14, 15] techniques have been reported as able to produce ohmic contacts and to suppress the ambipolar behavior of SBCNT-FETs. In the following, ideally ohmic doped source and drain contacts will be considered, according to the model presented in Sec. 1.2. The doping distribution will be considered uniform over the tube surface. The device channel is left intrinsic. The carrier charge density will be modeled as localized on the tube surface, thus leading to a volumetric charge

$$\rho(r, \theta, x) = \delta \left(r - \frac{d_{\text{CNT}}}{2} \right) \frac{q\left[p(x) - n(x) + N(x)\right]}{\pi d_{\text{CNT}}} \tag{2.11}$$

where (r, θ, x) represents the set of cylindrical coordinates, N is the net linear doping concentration in the source and drain regions, and p and n represent the linear hole and electron charge density, respectively. Similarly, the nanotube middle gap energy is calculated as the electrostatic potential at the tube shell

$$E_i(x) = -q\phi\left(r = \frac{d_{\text{CNT}}}{2}, x\right) \tag{2.12}$$

for every angular coordinate θ. Also the current flux is purely one-dimensional, flowing from source to drain, localized on the CNT surface and independent of the angular coordinate.

Due to the cylindrical symmetry of the device model, the simulation domain reduces to a 2-D region as reported in the right part of Fig. 2.5, where the different regions of the device are highlighted. The chosen architecture thus enables for a reduction of the complexity of the problem with a purely 2-D treatment of the electrostatic problem, coupled with the 1-D transport equation solved within the UMS approach.

As far as the electron-phonon interaction is concerned, phonon parameters are taken from [16]. In the latter an analysis of the impact of different phonon modes is reported. In our simulations, acoustic phonons and one longitudinal-optical phonon mode are included. The effect of self-heating and hot phonon effects, presented in [17] using a semi-classical model based on the Boltzmann transport equation, are neglected. The model still accurately describes all the main transport limiting mechanisms, being able to correctly predict the scaling trends in terms of off-current and sub-threshold slope, and to predict the qualitative behavior of the on-current ballisticity as well.

Depending on the doping configuration used for the source and drain contacts, two main classes of devices are identified. If the same doping species are used for both contacts, a device equivalent to standard silicon n- or p-MOSFETs is obtained, referred to as *n-i-n* CNT-FET (dually *p-i-p* for a p-type FET). A numerical characterization and analysis of performance of this device will be carried out in Sec. 2.4.

In the case of opposite doping species at the contacts, an interesting novel architecture of tunneling-FET is obtained. Following the same notation based on the doping distribution, the device will be named as *p-i-n* or *n-i-p*. Experimental evidence of tunneling devices has been reported in [18, 19]. The review of the transport mechanisms of this device and an evaluation of its scaling limits will be given in Sec. 2.5.

2.3 Non-parabolicity and band-to-band tunneling

The mathematical model presented in the previous Chapter has been formulated within the parabolic effective-mass approximation. The effective-mass approx-

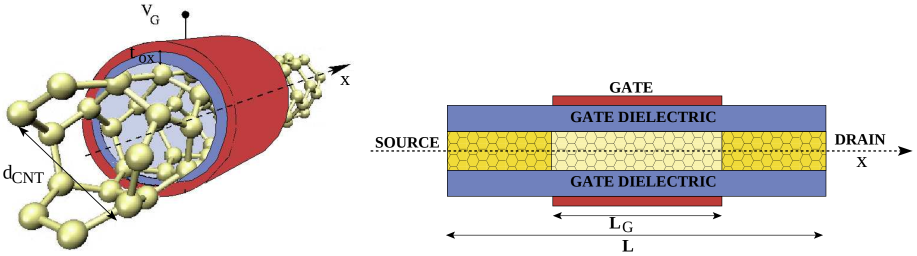

Figure 2.5: *(Left) Three-dimensional schematic representation of a GAA CNT-FET and (right) its projection on a plane along the transport direction. Air is supposed to fill the tube and the CNT surface is ideally treated as having zero-thickness.*

imation has been widely used in device simulation, in particular, in modeling silicon FETs, from standard single gate to more complex architectures. In order to have an accurate prediction of the effects of quantization on the silicon band gap for strongly confined electron gases, variations of the valley effective masses with respect to the bulk values have to be considered. At the same time, the lowering of the density of states associated to 2-D and 1-D systems have increased the importance of the deviation from a parabolic dispersion relation for degenerate systems. Limits of the assumption of parabolicity have been shown and model extensions with non-parabolic corrections have been proposed.

The same observations hold for the case of CNTs. Although in this work the modeling of metallic nanotubes is not treated, due to the clear impossibility of applying an effective-mass model, clear deviation from the parabolic approximations are evident also in the case of semiconducting CNTs. In addition, the band gap shown by technological feasible CNTs is narrower than the silicon one, thus the importance of quantum effects is increased. As reported for example in [18], CNT-FETs biased with small drain and gate voltages experience a drastic enhancement of the drain current due to the effect of band-to-band tunneling, further confirmed by simulation studies. An appropriate description of tunneling phenomena depends no longer on the treatment of the energy dispersion relation for traveling modes but on the accurate reproduction of the vanishing imaginary branches of the $E(k)$. The latter observation applies first of all to the description of the two subbands with smallest energy gap, namely the lowest conduction band and the highest valence one. An example of the complete band structure for a zigzag CNT is reported in Fig. 2.6, as calculated with a home-made solver using a nearest-neighbor TB model [20]. Due to the localization of the conduction and valence bands relative minimum or maximum at $k = 0$, the imaginary part of $E(k)$ lays completely in the $\Re\{k\} = 0$ plane.

In order to include the description of band-to-band tunneling phenomena in the transport model, a coupled solution of the carrier transport equation in con-

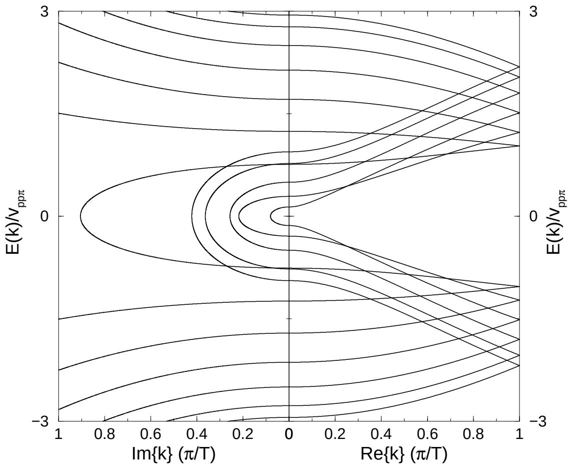

Figure 2.6: *Complete representation of the energy dispersion relation for a semiconducting CNT including both real and imaginary branches. The TB parameters are the same of the previous case.*

duction and valence band has been determined. Although 1-D subbands having different quantization indexes are treated separately, a unique equation is assumed for each couple of valence and conduction bands having the same index. A continuous transition is assumed for a carrier moving on a given subband couple extending the effective-mass Hamiltonian for the imaginary value of k, describing the vanishing states in the band gap. Due to the symmetric band structure, an effective mass and a density of states symmetric with respect to the midgap energy are assumed and associated to the transport description of electrons in conduction band or hole in valence band. As a self-consistent potential is imposed, the midgap or neutrality level $E_i(x)$ will be position dependent. Hence, the calculation of the total linear charge density in the device, according to the relations presented in the Sec. 1.2, reads out

$$\begin{aligned} Q(x_i) &= q\left(p(x_i) - n(x_i)\right) \\ &= \frac{2qi}{\Delta_x}\sum_b \sigma_b \left(\int_{E_i(x_i)}^{+\infty} \frac{\mathrm{d}E}{2\pi} G_b^{<}(i,i;E) - \int_{-\infty}^{E_i(x_i)} \frac{\mathrm{d}E}{2\pi} G_b^{>}(i,i;E) \right) \end{aligned} \tag{2.13}$$

where b is the quantization number spanning over all the conduction and valence band pairs, Δ_x is the discretization step along the transport direction, and σ_b accounts for the valley degeneracy. It is worth noting that the upper and lower integration extremes for the first and second term in the right hand side of (2.13) can be numerically reduced to a range of few tens of k_bT due to the presence of the Fermi functions causing the integrand to vanish. The calculation of the

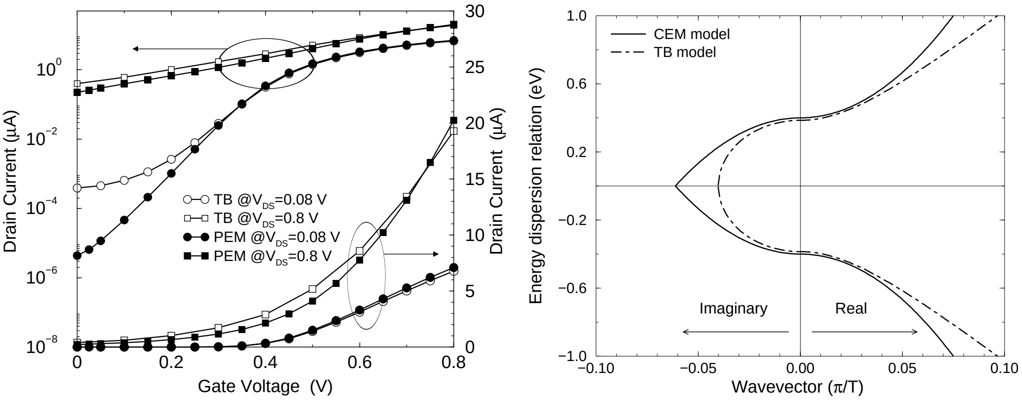

Figure 2.7: *(Left) Turn-on characteristics obtained with the CEM and TB model for $V_{\mathrm{DS}} = 0.08$ V and 0.8 V, in both logarithmic and linear scale. (Right) Energy dispersion relationships for the lowest conduction/valence subband pair of a $(13, 0)$ CNT calculated with the (dashed line) TB and (continuous line) CEM model, as a function of the normalized wavevector. The left part corresponds to energies within the gap.*

current, in the approximation of UMS, is

$$I = \frac{2q}{\hbar} \sum_{b} \sigma_b \int_{-\infty}^{+\infty} \frac{\mathrm{d}E}{2\pi} 2\Re \left\{ H_b\left(i, i+1; E\right) G_b^{<}\left(i+1, i; E\right) \right\} \ , \tag{2.14}$$

where H_b is the Hamiltonian for the b-th subband, and the other relations for the calculation of self-energy functions are left unchanged.

A first validation of the model was carried out with a direct comparison with the simulation results of an atomistic solver based on the single p_z-orbital TB model, as presented in [21]. The device is a (13,0) zigzag CNT with a diameter $d_{\mathrm{CNT}} = 1.0$ nm. The band gap and effective mass used in the presented simulator are taken from [22]. Like the band gap, the effective mass depends on the tube chirality. The channel, source, and drain lengths are fixed at 10 nm. No scattering mechanisms are included. Although small differences are present in the gate architecture, the choice of a thin gate oxide, $t_{\mathrm{ox}} = 1.5$ nm, ensures a similar electrostatic control for the two devices.

The comparison of the output current for $V_{\mathrm{D}S} = 0.08$ V and $V_{\mathrm{D}S} = 0.8$ V is reported in Fig. 2.7 (left), where the results obtained with the effective-mass approach are labeled as Constant Effective Mass (CEM). The two models provide a similar overall behavior, but remarkable differences are quite evident. The most striking one occurs in the sub-threshold region for low $V_{\mathrm{D}S}$, where the CEM model underestimates the current at $V_{GS} = 0$ V by up to almost two orders of magnitude. Particular care is therefore necessary when drawing conclusions about the sub-threshold behavior of the device in the frame of the CEM model.

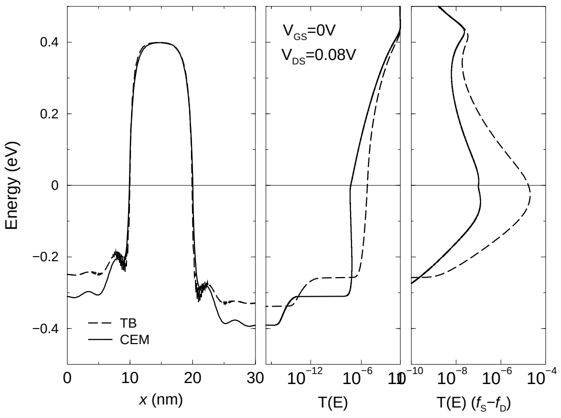

Figure 2.8: *(Left) Conduction-band edges as a function of x computed with the CEM and TB models for $V_{\mathrm{DS}} = 0.08$ V and $V_{\mathrm{GS}} = 0$ V. (Center) transmission coefficient $T(E)$ (horizontal axis). (Right) product of $T(E)$ with the difference of the source and drain Fermi functions $f_S - f_D$ (horizontal axis).*

The explanation for this large difference can be found in the different shapes of the band models near the midgap. As reported in Fig. 2.7 (right), showing the first complex subband calculated with the TB together with the one obtained by the CEM model, the two curves are in good agreement over a limited range of energies (about 100 meV) near the minimum of the conduction band (0.4 eV) and the maximum of the valence band (−0.4 eV). In the imaginary part of the diagram (band gap) a discrepancy is visible, with the largest difference at the midgap (0 eV), where the CEM model exhibits an unphysical discontinuous derivative. The divergence at large imaginary wave vectors for energies in the gap is reflected by an under-estimation of carrier tunneling probability and, consequently, of current.

The situation is further illustrated in Fig. 2.8, where the bottom-energy profiles as a function of x, the transmission probabilities $T(E)$ and the spectral current density, given by the product of $T(E)$ with the difference of the source and drain Fermi functions, f_S and f_D, are reported for the two models at $V_{\mathrm{DS}} = 0.08$ V and $V_{\mathrm{GS}} = 0$ V. The energy reference is the Fermi level in the source.

From the curves of the spectral density it appears that most of the current is due to electrons tunneling through the barrier at energies close to the midgap, where the two band models, and hence $T(E)$, primarily differ. The difference of two orders of magnitude in the current can then be explained by the difference observed in correspondence of the largest peaks in the $T(E) \times (f_S - f_D)$ plot. It is also to be noticed that the energy profiles are nearly identical in the channel, while discrepancies occur in the reservoirs where the electron gas is in a degenerate condition and the CEM approach forces a lower potential because of the lower density-of-states [2].

In order to cover the difference between the two models a correction to the

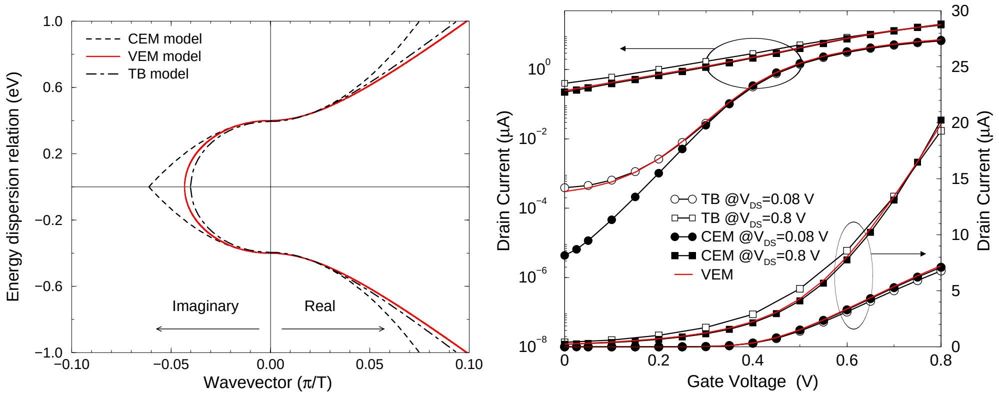

Figure 2.9: *(Left) Energy dispersion relation for the lowest conduction/valence subband pair of a* $(13, 0)$ *CNT calculated with the TB, the CEM, and (red solid line) the VEM model, as a function of the normalized wavevector. (Right) The same as in Fig. 2.7 but with the inclusion of the results obtained with the VEM model.*

standard parabolic model has been introduced. Following the extension of the effective-mass approach to a two-band scheme as proposed in [23], and subsequently applied to the analysis of FETs in [24], a non-parabolic energy dispersion relation $\varepsilon_b(k)$ is considered

$$\left(\varepsilon_b - \frac{E_{g_b}}{2}\right)\left(\frac{1}{2} + \frac{\varepsilon_b}{E_{g_b}}\right) = \frac{\hbar^2 k^2}{2m_b^\star}\,, \tag{2.15}$$

where $m_b^\star$ is the effective mass for the b-th subband. From (2.15) a parabolic Hamiltonian is derived with the introduction of a position-dependent effective mass

$$m_b(x;E) = \begin{cases} m_b^\star \left[1 + \frac{E - E_{c_b}(x)}{E_{g_b}}\right] & \text{if } E > E_i(x) \\ m_b^\star \left[1 + \frac{E_{v_b}(x) - E}{E_{g_b}}\right] & \text{if } E < E_i(x) \end{cases} \tag{2.16}$$

where E_{c_b} and E_{v_b} are the conduction and valence band-edges, respectively. As reported in Fig. 2.9 (left) the new energy dispersion relation, named Variable Effective Mass (VEM), is able to reproduce the results of the tight-binding model. The two curves are in good agreement over an extended range of energies, and, in the imaginary part of the diagram, the non-parabolic dispersion relation nicely fits the TB one.

The use of the VEM model in the transport analysis is then capable of covering the difference previously reported in the transfer characteristic for $V_{\mathrm{DS}} = 0.08$ V and low gate voltages. The new comparison is shown in Fig. 2.9 (right). The correction in the imaginary branch of the $E(k)$ is able to give the right prediction of the tunneling probability, with a correction of the current spectral density directly reflected in the resulting total current.

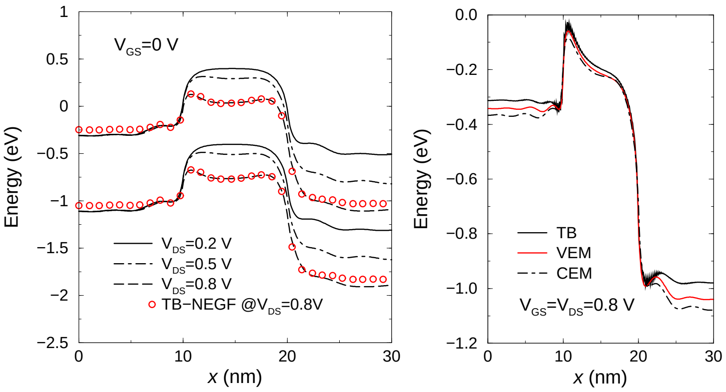

Figure 2.10: *(Left) Bottom of the conduction band and top of the valence band as a function of* x *computed at* $V_{\mathrm{GS}} = 0$ *V and different* V_{DS} *within the CEM model and the TB model (only* $V_{\mathrm{DS}} = 0.8$ *V). (Right) Comparison of the bottom of the conduction band as a function of* x *computed at* $V_{\mathrm{GS}} = V_{\mathrm{DS}} = 0.8$ *V within the three different models.*

Differences between the two models are in any case still present. An important feature of the turn-on characteristics is given by the device behavior for large drain biases. High currents are indeed reported in the sub-threshold regime by both models. The origin of this behavior can be investigated with the help of Fig. 2.10. Here, the bottom energies of the conduction band and the top energies of the valence band, computed from the CEM model for $V_{\mathrm{GS}} = 0$ V and different values of V_{DS}, are reported as a function of the coordinate along the CNT axis (x direction). As V_{DS} increases, a strong decrease of the source-to-channel energy barrier is observed, similar to the drain-induced-barrier-lowering (DIBL) in conventional MOS transistors, which is responsible of the large off-current at large V_{DS}. This effect was already noticed in [24], where it has been attributed to positive (hole) charge accumulation in the channel, due to electrons tunneling from the valence band in the channel region to the conduction band in the drain region, and for this reason also referred to as hole-induced-barrier-lowering (HIBL) [25]. In this condition, current mainly flows for thermionic emission over the top of the conduction band barrier. The difference between the TB and the VEM model, the latter lightly differing from the CEM one, can be thus ascribed to a different height of the energy potential barriers. This is the result of a self-consistent solution where differences in the electrostatic control and also in the modeling of the density-of-states at large kinetic energies are still present. The comparison of the energy profiles for the first subband as derived from the TB, CEM, and VEM model, is reported on the right side of Fig. 2.10 for a large drain voltage in saturation condition ($V_{\mathrm{DS}} = V_{\mathrm{DS}} = 0.8$ V). It clearly evidences the small differences in the band-edge in the channel, which directly reflect on different current values, and how the VEM model partially covers the

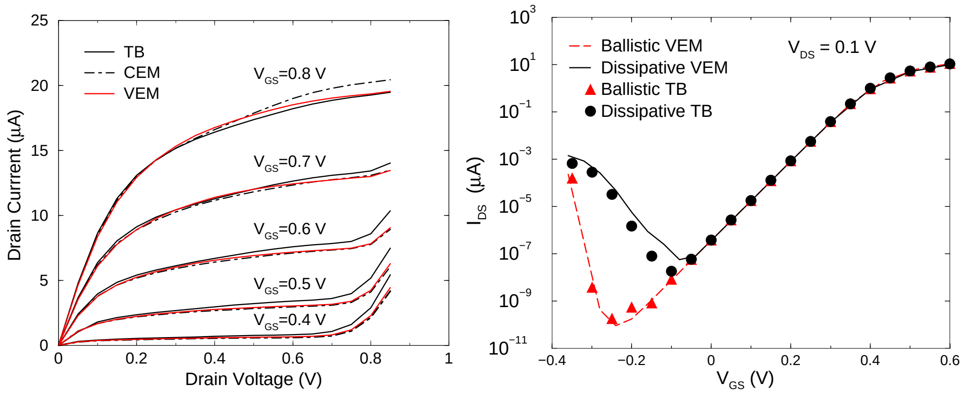

Figure 2.11: *(Left) Output characteristics of the CNT-FET obtained with the three models. (Right) Turn-on characteristics of an n-i-n* $(13, 0)$ *CNT-FET at* $V_{\mathrm{DS}} = 0.1$ *V. Lines: VEM model in the (dashed) ballistic and (solid) dissipative cases. Symbols: TB results from [28].*

differences with the TB model in the source and drain regions as well by imposing a larger density-of-states with respect to the CEM one. It is important in any case to highlight once more that the primary aim of the model, the inclusion of a qualitatively and quantitatively valid description of BTBT, is nicely fulfilled, giving an optimal compromise between accuracy and computational costs when compared with more complex models like the atomistic TB one.

The device output characteristics have been simulated as well. The simulation results are reported in Fig. 2.11 where the three different models are compared over a large range of gate and drain biases. An overall nice agreement of the VEM approach with the TB results is reported, reducing the difference on the on-current, previously justified as a result of a different density-of-states. For the smallest V_{GS} a sharp enhancement of the output current is predicted at large drain biases due to the HIBL. Additional degradation is found for V_{DS} larger than the band gap energy, due to the onset of a BTBT current path flowing from the valence band at the source side to the conduction band at the drain side, nullifying the gate control on current [26]. These output characteristics show a strong limitation on the maximum allowed supply voltage, that has to be kept lower than E_g or even as low as $E_g/2$ to achieve the maximum performances in terms of $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ current ratio.

Finally, the validation of the proposed model has been carried also in presence of scattering mechanisms. In Fig. 2.11 the transfer characteristics of a *n*-*i*-*n* device are reported compared with the TB results given by [28]. A very good agreement is found in the whole range of gate biases, in both cases of ballistic and dissipative transport. In particular, the onset of the ambipolar conduction for $V_{\mathrm{GS}} < 0$ V is well captured by the VEM model. This unexpected enhancement of the drain current is explained by a double band-to-band tunneling process between the conduction and valence band. In the ballistic case, the process becomes possible when the gate voltage is low enough to create a direct facing of the top of the valence band in the channel with the conduction band at the source and drain side. A current path connecting the drain and source reservoirs is thus created, avoiding the FET to be turned-off. It is interesting to note how the current increase for negative gate biases shows a slope steeper than when the device is turned-on. This is connected to the tunneling nature of the current flow, having no thermal limitations as compared to the thermionic emission. The sharp difference in the behavior between the ballistic and the dissipative case will be further investigated in the next section.

2.4 Performance evaluation of CNT-FETs

In this section an analysis of *n-i-n* CNT-FETs is presented. The device architecture is the same of that described in Sec. 2.2. The impact of acoustic and optical phonon scattering on the performance of CNT-FETs is investigated, showing the dependence on gate lengths, dielectric materials and chiralities. The analysis is also extended to devices with source/drain underlap. The effects induced by the electron-phonon scattering are addressed by comparing ballistic and dissipative quantum transport results. It has been found that the use of a high-κ dielectric improves the off-current limited by phonon-assisted band-to-band tunneling. In addition, the device scalability has been analyzed showing that with an oxide thickness of 1.5 nm, good performance metrics are obtained down to a gate length of 15 nm with SiO_2, and, on the other side, by the use of HfO_2 as gate dielectric, the gate length scaling can be extended down to 10 nm. The results confirm also that the calculation of the off-currents and delay times are strongly influenced by phonon scattering, whose inclusion in the transport model is fundamental for having realistic performance predictions.

2.4.1 Fixed gate length

In order to examine the impact of the phonon scattering on the electrical characteristics of scaled CNT-FETs, a first reference device is analyzed to understand the transport properties at different biases. The simulated FET has a $(13, 0)$ CNT, undoped channel and n-doped contacts ($N_D = 1.5 \times 10^7$ cm^{-1}). A cylindrical wrapped around SiO_2 layer with $t_{\text{ox}} = 1.5$ nm is used, and the gate length is fixed at $L_G = 20$ nm, for a total device length of 40 nm. The other electrical parameters are the same of the previous section. In Fig. 2.12 (left) the turn-on characteristics comparing the ballistic and the dissipative case are reported in logarithmic scale for $V_{\text{DS}} = 0.1$, 0.45 and 0.8 V. It is clearly shown that the effect of phonon scattering strongly depends on the applied drain voltage. The

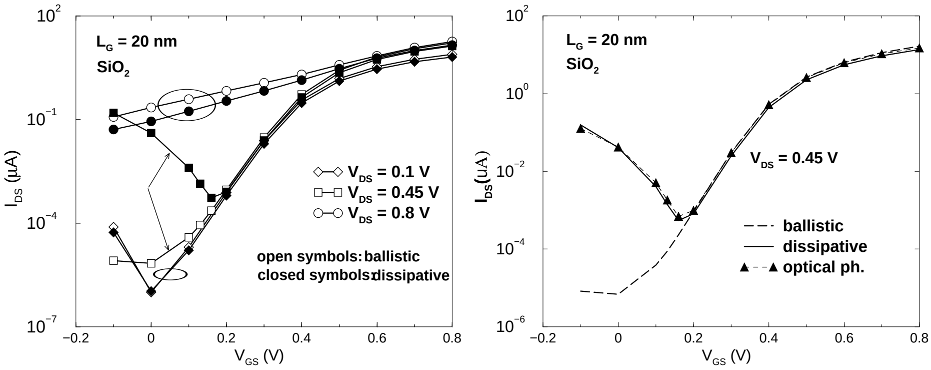

Figure 2.12: *(Left) Turn-on characteristic of a* 20*-nm* $(13, 0)$ *CNT-FET. Ballistic and dissipative calculations are reported for* $V_{\mathrm{DS}} = 0.1, 0.45$ *and* 0.8 *V. (Right) Ballistic, dissipative and only optical-phonon contribution at* $V_{\mathrm{DS}} = 0.45$ *V are isolated.*

explanation of the different curves is derived by observing the contour plots of the current density spectra in Fig. 2.13 at $V_{\mathrm{GS}} = 0.1$ V for the three different V_{DS} biases.

At low drain voltage, transport is mainly due to direct tunneling (DT) just below the channel barrier, and therefore the current is not influenced by phonons. At $V_{\mathrm{DS}} = 0.45$ V (Fig. 2.13, center) transport is dominated by optical-phonon-assisted band-to-band tunneling. Due to the additional phonon-assisted current fluxes within the channel band gap, the threshold of the ambipolar conduction is shifted to higher gate voltages with respect to the ballistic case. In this case a direct facing of the top of the valence band with the conduction band source is no longer requested, due to the coupling by phonon emission of carriers at energies with a relative difference of $\hbar\omega_{\mathrm{op}}$. In addition, the broadening of the density spectral function for the quasi-bond states in valence band, due to the electron-phonon interaction, smoothens the resonant behavior of the ambipolar conduction, as evidenced by the lower slope of the transfer characteristics (see also Fig. 2.11, right). This fundamental role played by optical phonons is confirmed by simulations reported in Fig. 2.12 (right), where their contribution to current has been isolated, showing no differences with the result obtained with the additional contribution of acoustic phonons. At high V_{DS} (Fig. 2.13, right), the electrostatic effect of the positive charge accumulation in the valence band, due to tunneling through the channel-to-drain interface, self-consistently lowers the source-to-channel barrier, inducing a high over-the-barrier current. The increase of the hole density produced by optical-phonon scattering is reflected on a different height of the barrier if compared with the ballistic case, inducing a lower I_{OFF} current [31].

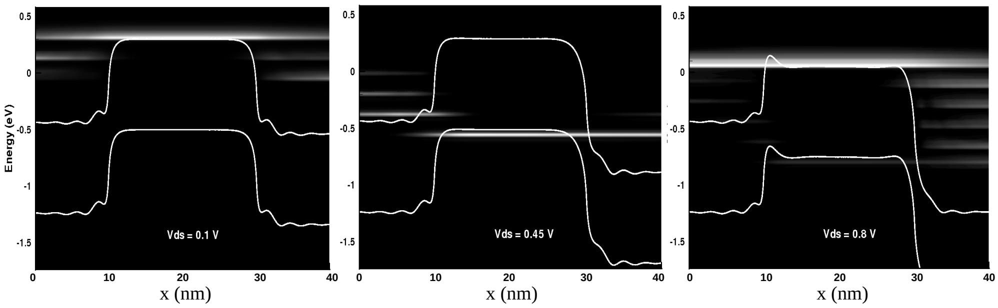

Figure 2.13: *Current spectral densities with phonon scattering of a* 20*-nm* $(13,0)$ *CNT-FET at* $V_{\mathrm{GS}}=0.1$ *V for* $V_{\mathrm{DS}}=0.1$ *(left),* 0.45 *(center) and* 0.8 *V (right).*

2.4.2 Effect of gate length scaling

The impact of the gate length scaling has been investigated for three different values of $L_G = 10$, 15, and 20 nm. The obtained turn-on characteristics are reported Fig. 2.14, for a drain bias of $V_{\mathrm{DS}} = 0.45$ V. At $L_G = 10$ nm, in the ballistic case, the DT contribution becomes high enough to increase the inverse sub-threshold slope (SS= $\partial V_{\mathrm{GS}}/\partial \log(I_{\mathrm{DS}})$), inducing a degradation from 68 mV/dec to 95 mV/dec at $V_{\mathrm{GS}} = 0.25$ V. When phonon scattering is turned on, the phonon-assisted BTBT clearly prevails for $V_{\mathrm{GS}} = 0$, and it is independent of L_G, being controlled by the shape of the conduction and valence band-edges at the source-to-channel junction rather than by the gate length. Due to the specification on the off-current level [27], the CNT-FETs will reach a lower scaling limit for $L_G = 15$ nm. The inset in Fig. 2.14 (left) shows the ballisticity, defined as the ratio between currents with and without scattering, $I_{\mathrm{scat}}/I_{\mathrm{bal}}$, as a function of the gate voltage in the above-threshold regime at the same drain voltage. As expected, the highest ballisticity is obtained with the smallest L_G, due to the reduced scattering effects.

2.4.3 Effect of changing gate dielectric

In Fig. 2.14 (right) the effects of phonon scattering when using HfO_2 as opposed to SiO_2 as dielectric has been investigated. In the analysis, t_{ox} is still fixed at 1.5 nm, thus the device with high-κ dielectric shows higher on-current and better sub-threshold slope due to the improved electrostatic control. The ballistic curves indicate that using high-κ material strongly increases also the BTBT current. This is a drawback of the higher gate capacitance. Due to the improved electrostatic control, the device with HfO_2 dielectric shows a higher barrier and hence a larger energy interval where BTBT can occur. On the contrary, in the dissipative case the phonon-assisted BTBT with HfO_2 is slightly less than with SiO_2. This is explained by considering that the phonon-assisted BTBT is more

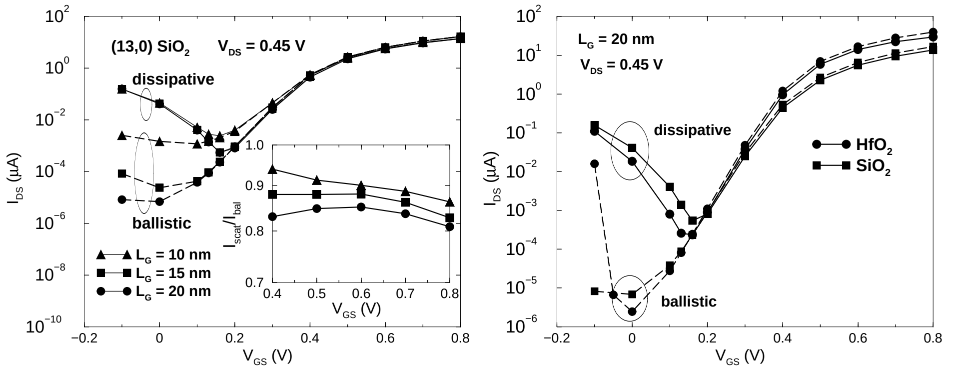

Figure 2.14: *(Left) Turn-on characteristic of a* $(13, 0)$ *CNT-FET at* $V_{\mathrm{DS}} = 0.45$ *V with three different gate lengths. Inset: Ballisticity vs. gate voltage at* $V_{\mathrm{DS}} = 0.45$ *V. (Right) Turn-on characteristics with different dielectric materials in both ballistic and dissipative case.*

sensitive to the tunneling distance the electrons have to travel in the gap rather than to the extension of the energy interval for which the conduction and valence bands face each other in the source-to-channel region. Indeed the larger capacitive fringing effect in the $\mathrm{HfO_2}$ case produces smoother barrier profiles at the channel-to-contact interfaces with respect to the $\mathrm{SiO_2}$ case and larger BTBT distances as well. The higher off-current for $\mathrm{SiO_2}$, together with the considerations for the on-currents, gives an overall better performance for the high-κ device.

It is also interesting to analyze the ballisticity in the case of standard and high-κ dielectric. In Fig. 2.15 a), the ballisticity curves are calculated separating the contributions of the different types of phonons. They indicate that the effect of phonons is more pronounced in the device with $\mathrm{HfO_2}$ with respect to the one with $\mathrm{SiO_2}$, in particular that of optical phonons at large V_{GS}. This can be ascribed to the more degenerate condition in the channel for a given gate voltage, that lowers the bottom energy of the conduction band with respect to the Fermi level, inducing a higher current, but, at the same time, a more favorable condition for optical phonon emission (the optical phonon energy is 180 meV).

2.4.4 Effect of CNT chirality

The phonon scattering is strongly influenced by the material properties: in particular, the electron-phonon coupling strength for a $(n, 0)$ CNT is found to be inversely dependent on n [16]. Thus, a similar analysis has been carried out on a CNT with chirality $(16, 0)$. The diameter is 1.25 nm and the energy gap is reduced to 0.65 eV. A high-κ gate dielectric is supposed. The $(16, 0)$ CNT-FET curves are compared with the $(13, 0)$ ones in the b) panel of Fig. 2.15. The band gap difference is directly reflected in a negative threshold voltage shift of about 150 meV. A small I_{ON} increase for a fixed gate overdrive is due to the different

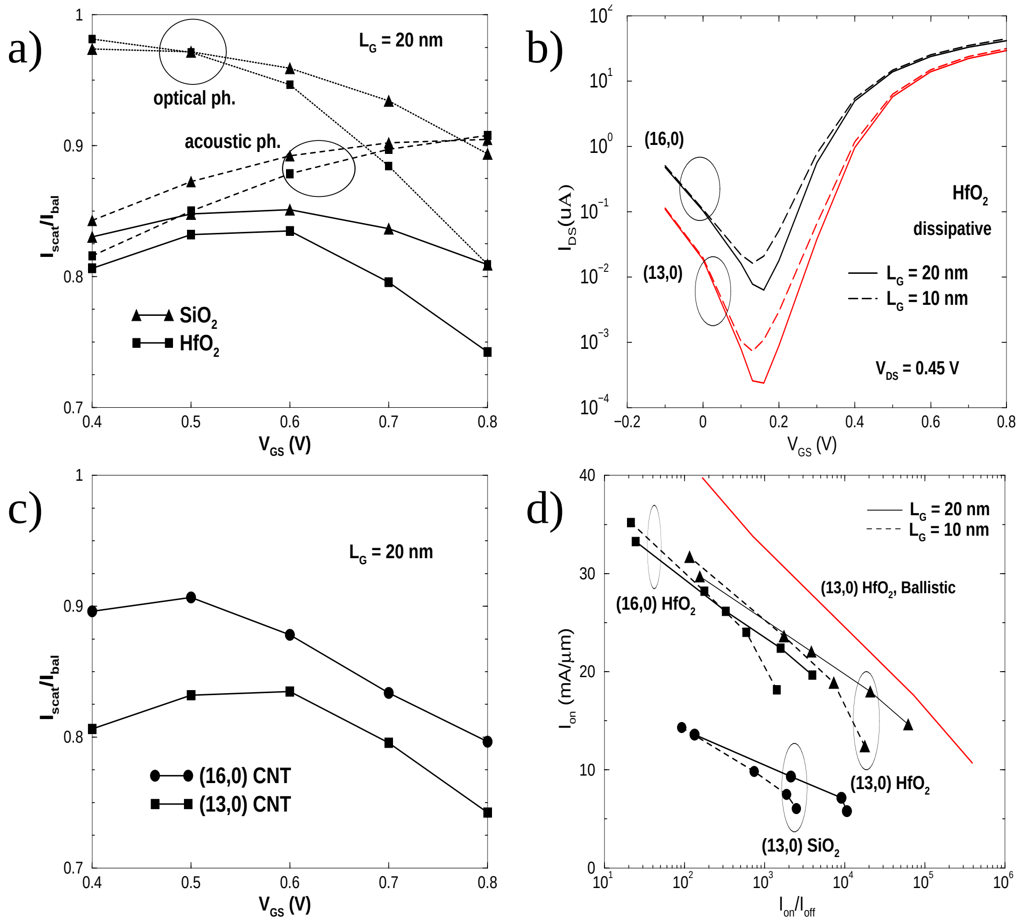

Figure 2.15: *a) Ballisticity as a function of the gate voltage for* $(13, 0)$ *CNT-FETs with* SiO_2 *and* HfO_2*. The different phonon contributions are isolated. b) Turn-on characteristics of the CNT-FETs with* HfO_2 *for two different chiralities at the same drain voltage. c) Ballisticity as a function of the gate voltage for* $(13, 0)$ *and* $(16, 0)$ *CNT-FETs with* HfO_2*. d) On-current as a function of the* I_{ON}/I_{OFF} *ratio for the different device configurations presented. A normalization of the current value with respect to the nanotube diameter is supposed.*

capacitive control as the tube diameter is increased. On the contrary, in the sub-threshold regime, the off-current is higher for the $(16, 0)$ CNT, as it is mainly affected by the reduction of the gap (that enhances the BTBT), in spite of the decrease of the phonon coupling constants. Indeed, in the above-threshold regime, the ballisticity shows a behavior in agreement with the relative strengths of the electron-phonon coupling constants, as reported in the c) panel of the figure.

All the previous results are finally collected in an analysis of the I_{ON} vs. I_{ON}/I_{OFF} (Fig. 2.15, d)) for the different CNT-FETs presented so far, as a further indication of the effect of phonon-scattering. The on-state of the device is defined

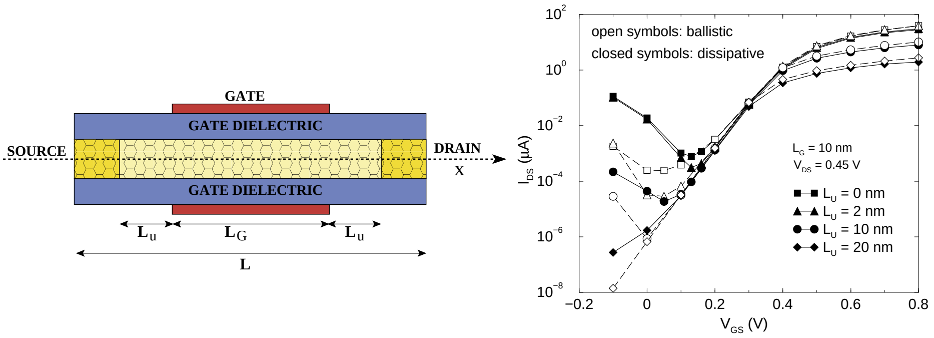

Figure 2.16: *(Left) Schematic view of the CNT-FET with underlapped source and drain. (Right) Turn-on characteristics of the underlap CNT-FET for different extensions L_U of the underlap region with fixed gate length ($L_G = 10$ nm) and total length ($L = 70$ nm).*

for a gate voltage of $V_{\mathrm{GS}}^{\mathrm{ON}} = V_{\mathrm{GS}}^{\mathrm{OFF}} + V_{\mathrm{DS}}$, where the drain voltage is fixed at $V_{\mathrm{DS}} = 0.45$ V. Consequently, the off-current is given as $I_{\mathrm{OFF}} = I(V_{\mathrm{GS}}^{\mathrm{OFF}})$. The plot is constructed as described in [32] by using a sliding $\left[V_{\mathrm{GS}}^{\mathrm{OFF}}, V_{\mathrm{GS}}^{\mathrm{ON}}\right]$ window. If $I_{\mathrm{ON}}/I_{\mathrm{OFF}} > 10^4$ is accepted as a constraint, the $(13, 0)$ CNT with HfO_2 can be scaled down to 10 nm, thus reducing the scaling limit of 5 nm with respect to the same device with SiO_2 dielectric. Beside, it appears that the ballistic calculation (red solid curve) largely overestimates the performance, confirming the necessity of a dissipative model for a realistic analysis.

2.4.5 CNT-FET with underlapped source/drain regions

In this section the importance of phonon scattering is examined for devices with underlapped source and drain regions, schematically illustrated in Fig. 2.16 (left). The intrinsic region extends for a length L_U on both sides of the gate. The total device and gate lengths are fixed at $L = 70$ nm and $L_G = 10$ nm, respectively. The dielectric material is HfO_2 with $t_{\mathrm{ox}} = 1.5$ nm, as above. The CNT chirality is $(13, 0)$. The effects of this architectural feature have been analyzed in [29, 30] on the basis of experimental data on CNT-FETs. Here, it refers to an alteration of the ideal device architecture ($L_U = 0$) considered so far.

The essential effects of introducing the underlap regions are illustrated by the turn-on characteristics of Fig. 2.16 (right). For increasing L_U, the strong beneficial reduction of the off-current is also accompanied by a detrimental reduction of the on-current. Once more, it is clearly seen that the behavior in the off-state, where BTBT is important, is not correctly interpreted by the ballistic model. In particular, for relatively small underlaps ($L_U = 2$ nm) the ballistic model largely overestimates the advantage of having the underlap, which instead becomes quite sizable for larger L_U. In the on-region the differences between the

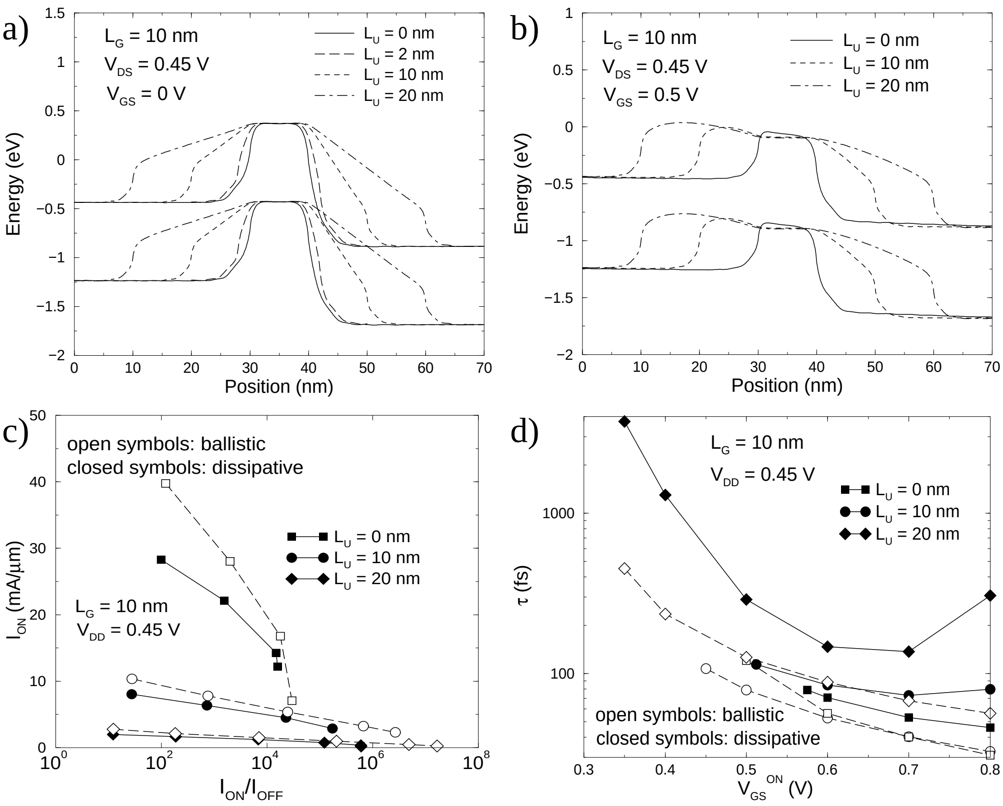

Figure 2.17: *Conduction and valence band edge profiles of the underlap CNT-FET for different extensions L_U of the underlap region in the a) off-bias and b) on-bias region. c) I_{ON} vs. $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ ratio and d) total delay vs. V_{GS} for the underlap CNT-FET for different L_U values. A normalization of the current value with respect to the nanotube diameter is supposed.*

two models are not as evident as in the off-region, and can be easily explained in terms of reduced ballisticity. To further illustrate the point, Fig. 2.17 a) reports the band-edge profiles for $V_{GS} = 0$ V in the dissipative case. The ballistic model gives substantially identical profiles, which are therefore not reported in the figure. For larger L_U the effective barrier width is increased, suppressing the DT component of the current. Also the BTBT current is strongly affected, due to the almost linear potential profiles in the underlap region.

Fig. 2.17 b) shows the band profiles in the on-bias condition ($V_{GS} = 0.5$ V). The current is limited by the barrier at the doped/intrinsic source junction, which prevents the device from being fully turned on. Besides, the barrier height increases somewhat with L_U, due to the decreasing electrostatic coupling with the gate, thus explaining the reduction of I_{ON}. Results on different cases are collected in the analysis of the $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ performance and reported in Fig. 2.17 c). It appears that the 10^4 ratio can be easily exceeded with a convenient underlap, at the cost of a significant reduction of the on-current. Similar results on the capability of improving the $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ ratio as a function of the gate/contact spacing have been shown in other simulation studies also in the case of Schottky barrier transistors [33].

The analysis of the CNT-FET performance is concluded with the calculation of the device delay time, defined as

$$\tau = \frac{1}{I_{\mathrm{ON}}} \int q \left[\rho_{\mathrm{ON}}(x) - \rho_{\mathrm{OFF}}(x) \right] \mathrm{d}x \ , \tag{2.17}$$

where $q \left[\rho_{\mathrm{ON}}(x) - \rho_{\mathrm{OFF}}(x) \right]$ is the total charge variation between the on- and off-states, and the integral is extended over the whole device. The parameter τ gives information about the delay in switching-on a load-transistor by a driver transistor of the same kind. The result is shown in the d) panel of Fig. 2.17 as a function of V_{GS} in the on-state. A significant degradation of τ is observed when a sizable underlap region is introduced, essentially due to the reduction of the on-current. As expected, the dissipative model predicts larger delay times, underlying the importance of taking into account the scattering mechanisms for accurate estimates of the speed performance. It is interesting also to observe that in the dissipative case with $L_U = 10$, and 20 nm there is an optimum choice of $V_{\mathrm{GS}}^{\mathrm{ON}}$ around $0.65 - 0.7$ V, in spite of the fact that I_{ON} increases with V_{GS}. This means that for larger values of the gate voltage the charge in the device increases faster than the on-current, which can be related with the importance of enhancement of optical phonon scattering when the channel is driven into strong degenerate conditions, as already observed in the previous section, due to the altered potential profile.

2.5 Performance evaluation of CNT Tunneling-FETs

It has been shown how aggressively scaled CNT-FETs suffer from a large increase of the sub-threshold current at large drain biases due to the charge pile-up effect, limiting the range of applicable voltages. In order to overcome this limitation, *p-i-n* CNT-FETs operating in the band-to-band tunneling regime have been proposed [18]. This type of device is based on concepts already extensively explored for both silicon and III-V materials. Sub 60 mV/dec inverse sub-threshold slopes have been reported by both experimental and theoretical analysis, with the advantage of very promising on/off current ratios. Here, the interest is focused in performing a computational study of the *p-i-n* CNT-FETs to investigate their properties when pushed to the extreme scaling limits. Taking as a reference the specifications for the double-gate (DG) MOSFET at the ITRS Low Operational Power technology node LOP32 [27], a $(13, 0)$ *p-i-n* CNT-FET has been preliminary optimized choosing proper source and drain concentrations and gate-to-channel overlap. Then the effects of changing the oxide thickness, power

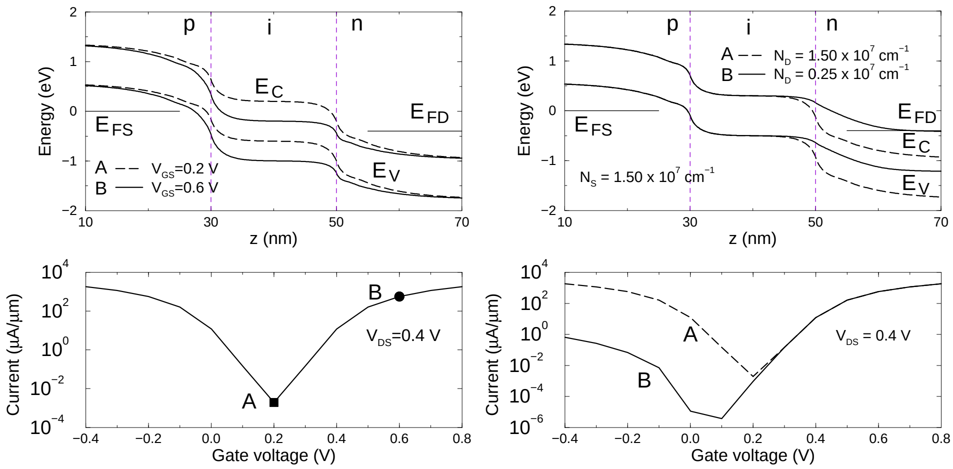

Figure 2.18: *(Left) p-i-n CNT-FET with symmetrically doped source and drain regions. Top panel: Lowest conduction (E_C) and highest valence (E_V) band-edges as a function of the longitudinal coordinate for the two bias conditions corresponding to points A and B of the bottom. The source and drain Fermi levels are also shown. Bottom panel: Drain current normalized to the CNT diameter vs. gate voltage for $V_{\mathrm{DS}} = 0.4$ V. (Right) Transport in symmetrically (label A) and asymmetrically (label B) devices at $V_{\mathrm{DS}} = 0.4$ V. N_S and N_D are the doping concentrations in the source and drain regions, respectively. Top panel: Conduction (E_C) and valence (E_V) band-edges as a function of the longitudinal coordinate for $V_{\mathrm{GS}} = 0.1$ V. Bottom panel: Drain currents normalized to the CNT diameter vs. gate voltage.*

supply and gate length have been systematically analyzed with the aim of finding the scaling limits and an optimal parameter set, using, as the key-reference, the off-current specified for the LOP DG-MOSFETs. The physical model is the same as described above, accounting for non-parabolicity through the use of an energy-dependent effective mass. The analysis will be mainly carried out in dissipative regime.

2.5.1 Fundamental transport properties

The reference architecture for the device is the same previously presented with a GAA gate configuration, intrinsic channel, and doped source and drain contacts. For the *p-i-n* device, the doping species are different for the two contacts, with an acceptor-doped source and a donor-doped drain contact. If band-to-band tunneling is limiting the performance of conventional *n-i-n* CNT-FET, this feature is turned into advantage in the *p-i-n* CNT-FETs, where the BTBT is the basic injection mechanism.

Current is controlled by the gate through the modulation of the tunneling barriers at both the source-to-channel and channel-to-drain junctions as illus-

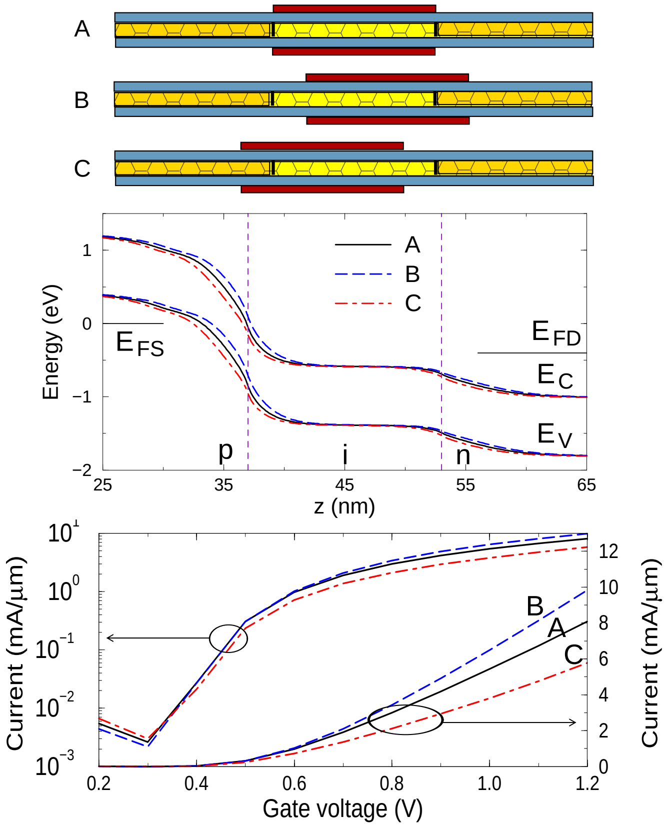

Figure 2.19: *(Top) Effect of gate-to-channel misalignments. Case A: aligned gate. Case B:* 1*-nm source underlap and* 1*-nm drain overlap condition. Case C:* 1*-nm source overlap and* 1*-nm drain underlap condition. (Center) Conduction and valence band-edges as a function of the longitudinal coordinate for* $V_{\rm DS} = 0.4$ *V and* $V_{\rm GS} = 1.0$ *V. (Bottom) Drain currents normalized to the CNT diameter vs. gate voltage for* $V_{\rm DS} = 0.4$ *V.*

trated in the top-left panel of Fig. 2.18, where the subband profiles of the first valence and conduction bands are reported for an off- and an on-biasing condition. In the bottom-left panel, an example of the turn-on characteristics is given in the ballistic case. Due to the equivalent gate modulation of both the tunneling barriers, the device is intrinsically ambipolar and, if symmetric source and drain doping concentrations are used, the turn-on characteristics are symmetric with respect to the $V_{\rm GS} = V_{\rm DS} = 2$ bias condition [34], in which the tunneling barrier at the source chemical potential energy level extends for the entire gate length. The most attractive feature of this device is given by its sub-threshold behavior. Since current is controlled by BTBT, the SS is not thermally limited and values lower than 60 mV/dec can be easily reached at room temperature, allowing for good $I_{\rm ON}/I_{\rm OFF}$ ratios and gate delays even at low supply voltages. Therefore this type of devices is considered more suitable for LOP rather than

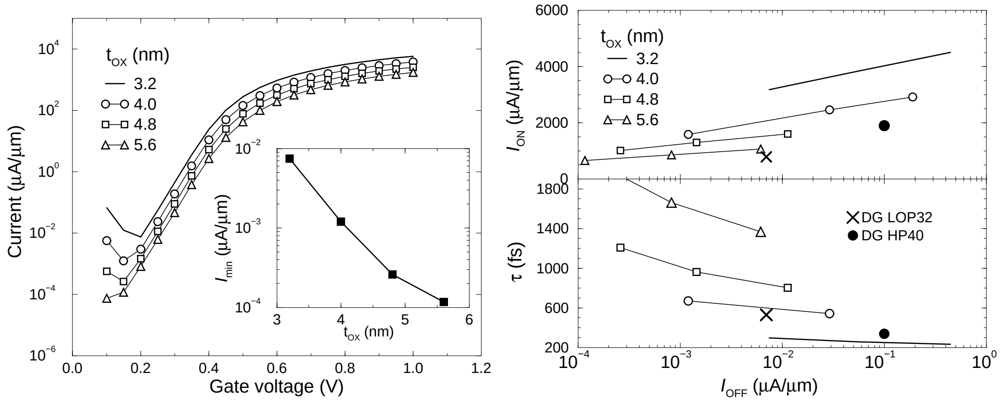

Figure 2.20: *(Left) Turn-on characteristics of a* $(13,0)$ *p-i-n CNT-FET with* $L_G = 16$ *nm and different oxide thickness at* $V_{\rm DS} = V_{\rm DD} = 0.6$ *V. Inset: Minimum off-current* $I_{\min}$ *vs.* $t_{\rm ox}$. *(Right) Device performance. Top panel:* $I_{\rm ON}$ *vs.* $I_{\rm OFF}$. *Bottom: gate-delay* τ *vs.* $I_{\rm OFF}$. *The points correspond to the DG-MOSFET performance of the LOP32 and HP40 ITRS technology nodes.*

for high-performance (HP) applications.

The ambipolar behavior is surely the most limiting feature of the device, naturally degrading the on- over off-current ratio. For this reason a preliminary optimization has been performed to make the tunneling process in the source-to-channel junction more favorable with respect to the drain-to-channel one, and finally obtaining a n-channel behavior, with increasing current for positive gate voltages. Asymmetric doping concentrations are used in the source and drain regions, with the acceptor concentration N_S in the source higher than that of the donors N_D in the drain. As seen in Fig. 2.18, top-right panel, the higher N_D, the steeper the potential profile at the drain-to-channel junction. A similar behavior can be observed for N_S. Hence, the condition $N_S > N_D$ creates a BTBT barrier at the drain junction wider than that at the source, causing asymmetric tunneling processes. This implies a shift of the current minimum to $V_{\rm GS} < V_{\rm DS}/2$ (Fig. 2.18, bottom-right). The choice of the optimal doping concentrations has been driven by the following considerations. In order to reach the largest $I_{\rm ON}$, $N_S = 1.5 \times 10^7$ cm^{-1} has been chosen, being a realistic concentration in terms of ratio between dopant ions and carbon atoms per unit length. To reach the lowest $I_{\rm OFF}$, a low N_D needs to be used in the drain. In this case, the lower boundary is dictated by the need of preserving a neutral contact region for a good electrostatic control. $N_D = 0.25 \times 10^7$ cm^{-1} has been chosen.

An additional architectural choice useful to emphasize the n-type transport is the use of a metal gate not aligned with respect to the intrinsic channel. A thorough analysis led to the conclusion that an overlap condition for the gate electrode at the source side would lead to a detrimental lowering of $I_{\rm ON}$. On

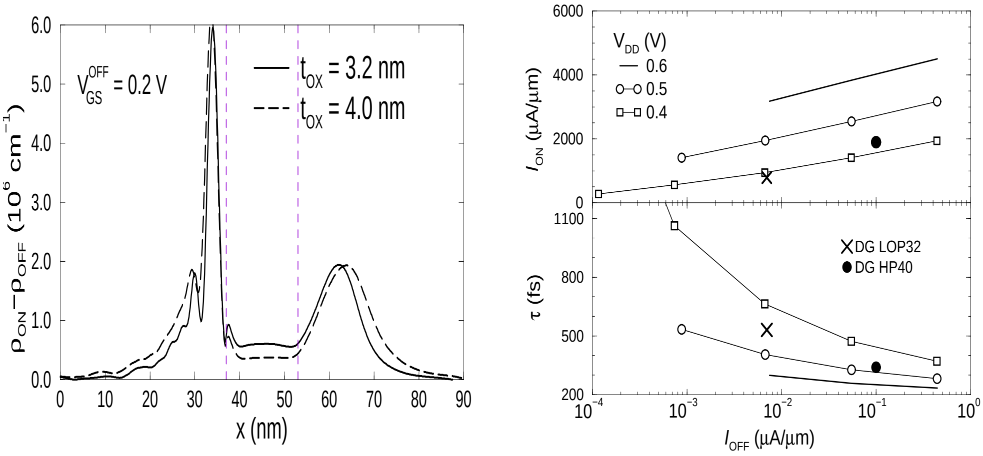

Figure 2.21: *(Left) Variation of the linear carriers density between the on- and off- state as a function of the longitudinal coordinate in a* $(13,0)$ *p-i-n CNT-FET with* $L_G = 16$ *nm for two different* t_{ox}*. The vertical dashed lines indicate the gate edges. (Right) Device performance for different supply voltages. Top panel:* I_{ON} *vs.* I_{OFF}*. Bottom: gate-delay* τ *vs.* I_{OFF}*. The points correspond to the DG-MOSFET performance of the LOP32 and HP40 ITRS technology nodes.*

the contrary, an underlap condition at the source side gives a steeper potential transition, rising I_{ON} and SS. An optimal underlap value exists which is a function of the oxide thickness, t_{ox}. The same considerations apply to the gate overlap/underlap at the drain side: a small amount of overlap is useful to smooth out the potential transition, the upper limit being set by the necessity of avoiding excessive parasitic capacitances. The effects on the potential energy profiles and turn-on characteristics for both underlap and overlap conditions are shown in Fig. 2.19, for the corresponding underlap/overlap conditions schematically presented in the top part of the figure. A $1-$nm underlap at the source and a $1-$nm overlap at the drain have been chosen, since they are close to the optimum for different values of t_{ox}.

2.5.2 Oxide thickness and voltage variations

Considering the optimized device presented above, with a gate length fixed at 16 nm and a high-κ oxide material, the effect of the oxide thickness scaling had been first considered. In Fig. 2.20 the turn-on characteristics are shown for different t_{ox} in the range $3.2 - 5.6$ nm for $V_{\mathrm{DS}} = V_{\mathrm{DD}} = 0.6$ V, with the aim of analyzing the trade-off between device performance and current leakage. Currents are normalized with respect to the CNT diameter. The minimum achievable current I_{min} versus t_{ox} is reported in the inset, highlighting a sharp dependence. The current variations are directly connected to the change of potential profiles

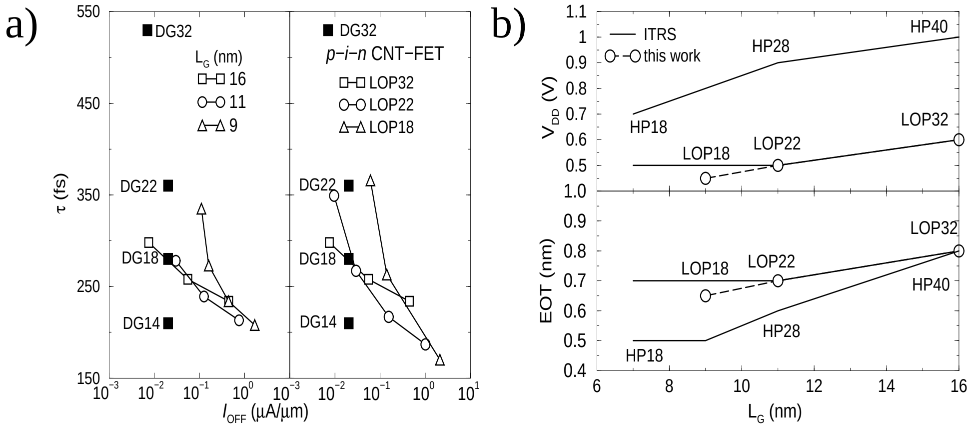

Figure 2.22: *a)* τ *vs.* I_{OFF} *performance of a* $(13,0)$ *p-i-n CNT-FET. Left panel: EOT=* 0.8 *nm* ($t_{\mathrm{ox}} = 3.2$ *nm) and* $V_{\mathrm{DD}} = 0.6$ *V fir different gate lengths. Right panel: EOT,* V_{DD} *and* L_G *are scaled according to the dashed line reported in the right part of the figure. The points correspond to the projected performance of the DG-MOSFETs of the ITRS LOP nodes. b) Scaling rules for (top panel) the supply voltage and (bottom panel) EOT for the ITRS DG-MOSFET HP and LOP technology nodes (solid lines). The scaling rules adopted in this work for the LOP devices are also shown (dashed lines with circles).*

of the tunneling junctions. For decreasing t_{ox}, thinner and sharper barriers are created at the junctions due to the changes in the electrostatic coupling between gate and doped regions, causing an increase of both I_{OFF} and I_{ON}. The resulting I_{ON} and gate-delay performance metrics are reported in the right of Fig. 2.20 as a function of I_{OFF}, together with the corresponding points for the DG-MOSFET at the LOP32 technology nodes. The gate-delay plot is realized by following the procedure already described above. A clear advantage in the $I_{\mathrm{ON}}/I_{\mathrm{OFF}}$ ratio with respect to the DG-MOSFET LOP32 is found for all t_{ox}, which is due to the low SS of the BTBT device (SS= 55 mV/dec, derived from the turn-on characteristics presented in figure Fig. 2.20, left). The performance of the LOP32 can be thus reached with a relaxed oxide thickness, with advantage on gate leakage currents (not estimated here). As far as τ versus I_{OFF} is concerned, the LOP32 requirements are satisfied for a range of t_{ox} between 3.2 and slightly less than 4 nm. For instance, for $t_{\mathrm{ox}} = 3.2$ nm the LOP32 off-current limit is reached with an almost 50% gain in gate-delay. For completeness the points for the DG-MOSFET HP40 node, characterized by the same $L_G = 16$ nm and EOT= 0.8 nm, but with $V_{\mathrm{DD}} = 1$ V, are reported in Fig. 2.20 (right). By analyzing both I_{ON} and τ, the *p-i-n* device clearly outperforms the HP40 DG-MOSFET as well.

To deeply understand the behavior of the *p-i-n* device, the dependence of τ on t_{ox} has been further investigated by analyzing the device charge variation. Fig. 2.21 (left) shows the total charge variation along the device for two different

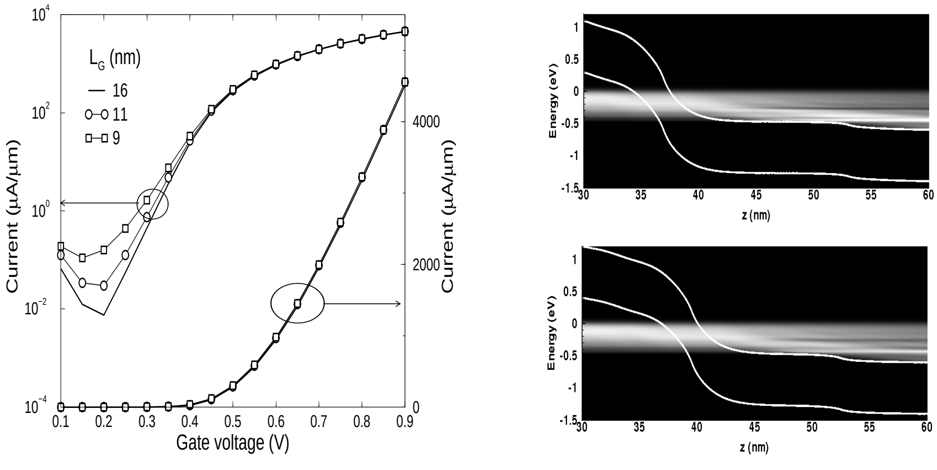

Figure 2.23: *(Left) Turn-on characteristics with EOT=* 0.8 *nm and* $V_{\mathrm{DD}} = 0.6$ *V for different gate lengths* L_G*. (Right) Spectral current density at* $V_{\mathrm{DS}} = V_{\mathrm{DD}} = 0.6$ *V and* $V_{\mathrm{GS}} = 0.9$ *V for two different gate lengths: (top panel)* $L_G = 16$ *nm and (bottom panel)* $L_G = 11$ *nm.*

values of the oxide thickness. It is clearly shown that the largest charge variation occurs outside the channel region, due to the capacitive gate-to-source and gate-to-drain fringing effects. This behavior is opposed to what is expected in a conventional MOSFET, where the largest variation takes place in the channel. By augmenting the oxide thickness, due to the increased relative importance of the fringing effects, the charge variation in the source and drain regions is even more pronounced. The overall effect is that the integrated charge variation increases with t_{ox}, and explains why the gate delay exhibits such a sharp degradation with increasing t_{ox}.

A similar analysis has been conducted considering the supply voltage scaling, while the gate length and the oxide thickness are kept fixed at 16 and 3.2 nm, respectively. Three values of drain voltage have been considered, $V_{\mathrm{DS}} = V_{\mathrm{DD}} = 0.6, 0.5$, and 0.4 V, focusing on the trade-off between device performance and power saving. Accordingly to previous results [34], a decrease of the I_{min} is found decreasing the power supply. At the same time, a similar degradation of the I_{ON} current is experienced. In the right part of Fig. 2.21 the on-current and gate-delay values are reported as a function of the off-current for the different V_{DD}. It is shown that there is a performance advantage with respect to the ITRS reference device (DG LOP32) even when the supply voltage is somewhat reduced. When comparing the *p-i-n* device with the HP40 DG-MOSFET, the advantage persists also for V_{DD} scaled down to 0.5 V.

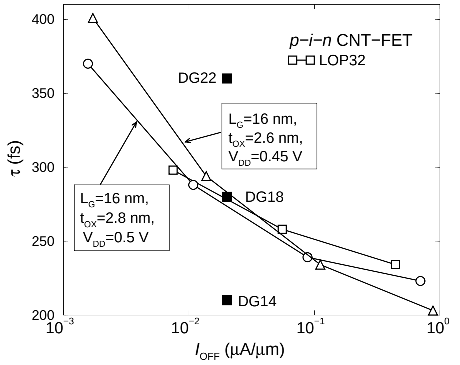

Figure 2.24: *τ vs. I_{OFF} performance of $(13, 0)$ p-i-n CNT-FETs with $L_G = 16$ nm. Open squares (LOP32): EOT= 0.8 nm ($t_{\text{ox}} = 3.2$ nm) and $V_{\text{DD}} = 0.6$ V. Circles: EOT= 0.7 nm ($t_{\text{ox}} = 2.8$ nm) and $V_{\text{DD}} = 0.5$ V. Triangles: EOT $= 0.65$ nm ($t_{\text{ox}} = 2.6$ nm) and $V_{\text{DD}} = 0.45$ V. Filled squares: projected performance of the DG-MOSFETs of the ITRS LOP node.*

2.5.3 Scaling trends

The final aim of the proposed computational study is to investigate the scaling trend of the *p-i-n* CNT-FETs focusing on LOP applications, for which, according to the above considerations, the maximum advantage is expected. Since the *p-i-n* CNT-FETs are non conventional devices, as the current injection mechanism is due to BTBT, first of all it is necessary to understand the impact of scaling L_G on the device characteristics. To this purpose, τ versus I_{OFF} is reported in left panel of Fig. 2.22 a) for different gate lengths. The L_G values are chosen in accordance with the specifications for the ITRS nodes LOP32, LOP22 and LOP18, but the same EOT= 0.8 nm and supply voltage $V_{\text{DD}} = 0.6$ V of the DG-MOSFET LOP32 node are used. The projected performance of the DG-MOSFET for the different nodes is also shown. It is seen that scaling L_G down to 11 nm leads to a modest reduction of τ for large I_{OFF}, and it is even detrimental at the DG18 off-current limit. At $L_G = 9$ nm the degradation of the gate delay is quite large for any acceptable value of I_{OFF}. The explanation of such behavior is threefold. Scaling L_G leads to i) a sub-threshold current increase (see Fig. 2.23 (left), logarithmic scale), due to the direct source-to-drain tunneling that considerably worsen SS and is particularly severe in short channel CNTs; ii) only a modest reduction of the total capacitance which is mainly due to the source/drain fringing contributions as explained above; iii) no increase in the on-current (see Fig. 2.23 (left), linear scale) due to the BTBT nature of the transport and not affected by an improved ballisticity. To better explain the latter aspect, the current spectral densities in an on-state bias are reported in Fig. 2.23 for the CNT-FETs with $L_G = 16$, and 11 nm. The presence of phonon scattering in the channel is clearly visible in both spectra. However, electrons injected from the source and scatter-

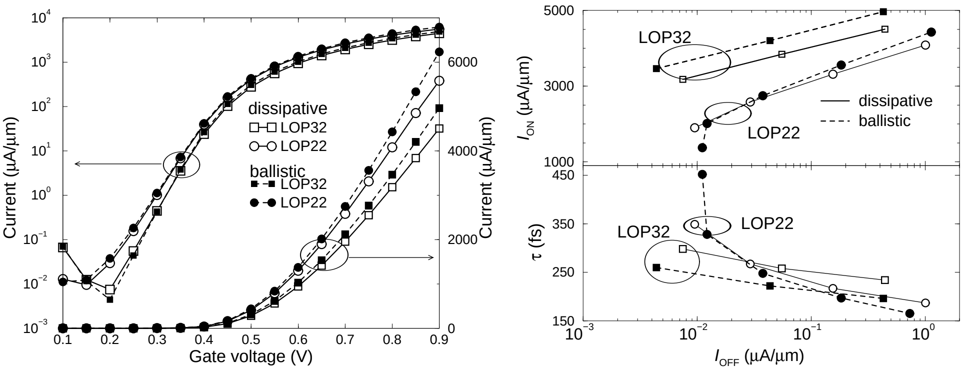

Figure 2.25: *(Left) Turn-on characteristics for the* $(13, 0)$ *pi-n CNT-FETs scaled according to the LOP32 and LOP22 rules of Fig. 2.22, computed with both ballistic and dissipative models at* $V_{DS} = V_{DD}$. *(Right) Top panel:* I_{ON} *vs.* I_{OFF} *performance; bottom panel:* τ *vs.* I_{OFF} *performance, both computed accordingly the sane scaling rules in the ballistic and dissipative case.*

ing in the channel have a low probability of being back-scattered into the source due to the energy selectivity of the BTBT. Therefore the current is affected by phonon scattering occurring mainly in the source region and source-to-channel junction, and is therefore independent on gate length.

The above analysis has been repeated for the same gate lengths but applying the scaling rules for EOT and V_{DD} extracted from ITRS for the LOP nodes and reported in Fig. 2.22 b). The τ versus I_{OFF} performance are shown in Fig. 2.22 a), right panel. As expected, a slight performance improvement is obtained with respect to the previous analysis (left panel of the same figure). Anyway, there is no advantage in scaling to the LOP22 node and a strong disadvantage in moving to the LOP18 node. From the above analysis it is therefore possible to conclude that the optimum gate length is around 16 nm, and that with such L_G the performance of the LOP18 DG-MOSFET is met. In an attempt to define the ultimate optimum device, the L_G =16 nm CNT-FET has been simulated with EOT and V_{DD} scaled in accordance to a revised trend, labeled as “this work” in Fig. 2.22 b). The result is shown in Fig. 2.24. It is seen that, for the specified I_{OFF} limit, changing EOT and the supply voltage does not lead to any sizable advantage in τ, but the reduction of V_{DD} allows for a proportional power saving.

As previously stated, all the presented results have been extracted through an analysis in dissipative condition. It remains in any case interesting to present some remarks on the importance of the electron-phonon scattering when dealing with the scaling trend of *p-i-n* CNT-FETs. The question arises because the neglect of phonon scattering would speed up the computation times considerably. The turn-on characteristics for the CNT-FETs scaled according to the LOP32 and LOP22 ITRS rules have been simulated again turning phonon scattering off (ballistic transport). The results are compared with the dissipative case in Fig. 2.25. In the longer channel device ($L_G = 16$ nm) phonon scattering slightly increases I_{min} due to its effect on BTBT (see [16], and results presented in the previous section), and decreases I_{ON} by almost 10%. In the LOP22 device the effect on I_{ON} is quantitatively similar, in spite of the shorter gate length ($L_G = 11$ nm), for the reasons explained in the above result analysis. On the contrary, I_{min} in the LOP22 is decreased by the phonon scattering: the sub-threshold current is dominated by direct source-to-drain tunneling, which gets reduced by the electron-phonon interactions breaking the electron phase-coherence. In Fig. 2.25, right, the I_{ON} versus I_{OFF} and τ versus I_{OFF} performance are reported for the same devices considered above. The purely ballistic analysis would lead to the same conclusion reached in the previous section, namely the lack of advantages in scaling down the gate length beyond $L_G = 16$ nm. However, for a quantitative current estimation, the inclusion of phonon scattering cannot be neglected even at such short gate lengths.

2.6 Summary

In this Chapter a full quantum description of the transport problem has been applied to the analysis and optimization of CNT-FETs. After a brief overview on the electronic properties of CNTs, an extension of the parabolic effective-mass approximation is presented which includes the effects of the non-parabolicity of the energy dispersion relation and the band-to-band tunneling. A validation of the model has been given in both ballistic and dissipative transport regime through a comparison with a more physics-based model (tight-binding). The presented model, extending similar proposals presented in the literature, is found to be a valid alternative to the full tight-binding approach due to its remarkable computational time advantage and constitutes one of the main results presented in this work. Two different device architectures have been deeply investigated: a MOSFET-like FET and a tunneling-FET. The limitations of the scaling process, the dependence on the architectural choices, and the role played by the electron-phonon interaction have been the subject of a thorough physical analysis. In particular, the advantages of using the *p-i-n* tunnel-FET architecture for low operational power applications has been discussed.

Bibliography

[1] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical properties of carbon nanotubes*, Imperial College Press, London, 1998. [cited at p. 32, 34, 35, 36, 83, 87]

[2] Jing Guo, *Carbon Naotube Electronics: Modeling, Physics, and Applications*, Ph.D. Thesis, Purdue University, August, 2004. [cited at p. 34, 42]

[3] P. L. McEuen, M. S. Fuhrer, and H. K. Park, "Single-walled carbon nanotube electronics," *IEEE Trans. on Nanotech.*, vol. 1, pp. 78–85, 2002. [cited at p. 34]

[4] V. Zólyomi and J. Kürti, "First-principles calculations for the electronic band structures of small diameter single-wall carbon nanotubes" *Phys. Rev. B*, vol. 70, p. 085403, 2004. [cited at p. 36]

[5] S. Iijima and T. Hichihashi, "Single-shell carbon nanotubes of 1-nm diameter" *Nature*, vol. 363, pp. 603–605, 1993. [cited at p. 37]

[6] J. W. G. Wildöer, L. C. Venema, A. G. Rinzler, R. E Smalley, and C. Dekker "Electronic structure of atomically resolved carbon nanotubes" *Nature*, vol. 391, pp. 59–62, 1998. [cited at p. 37]

[7] C. Zhang *et al.*, "Selective etching of Metallic Carbon Nanotubes by Gas-Phase Reaction" *Science*, vol. 314, pp. 974–977, 2006. [cited at p. 37]

[8] J. W. Odom, J.-L. Huang, P. Kim, and C. M. Lieber "Atomic structure and electronic properties of single-walled carbon nanotubes" *Nature*, vol. 391, pp. 56–64, 1998. [cited at p. 37]

[9] J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, and P. Avouris, "Field-Modulated Carrier Transport in Carbon Nanotube Transistors" *Phys. Rev. Lett.*, vol. 89, no. 12, p. 126801, 2002. [cited at p. 37]

[10] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, and P. Avouris, “Carbon Nanotubes as Schottky Barrier Transistors” *Phys. Rev. Lett.*, vol. 89, no. 10, p. 106801, 2002. [cited at p. 37]

[11] A. Javey *et al.* “Self-aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays”, *Nano Lett.*, vol. 4, no. 7, pp. 319–322, 2004. [cited at p. 37]

[12] Y.-M. Lin, J. Appenzeller, J. Knoch, and P. Avouris, “High-Performance Carbon Nanotube Field-Effect Transistor With Tunable Polarities”, *IEEE Trans. on Nanotech.*, vol. 4, no. 5, p. 481, 2005. [cited at p. 37]

[13] J. Chen, C. Klinke, A. Afzali, K. Chan, and P. Avouris, “Self-aligned carbon nanotube transistors with novel chemical doping”, *IEDM Tech. Dig.*, pp. 695-698, 2004. [cited at p. 37]

[14] A. Javey *et al.* “High Performance n-Type Carbon Nanotube Field-Effect Transistors with Chemically Doped Contacts”, *Nano Lett.*, vol. 5, no. 2, pp. 345–348, 2005. [cited at p. 37]

[15] Y. Lu, S. Bangsaruntip, X. Wang, L. Zhang, Y. Nishi, and H. Dai , “DNA Functionalization of Carbon Nanotubes for ultrathin Atomic Layer Deposition of High-κ Dielectrics for nanotube Transistors with 60 mV/Decade switching”, *J. Am. Chem. Soc.*, vol. 128, pp. 3518–3519, 2006. [cited at p. 37]

[16] S.O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, “Ballisticity of nanotube field-effect transistors: Role of phonon energy and gate bias”, *Appl. Phys. Lett.*, vol. 89, p. 023125, 2006. [cited at p. 38, 49, 62]

[17] S. Hasan, M. A. Alam, and M. S. Lundstrom, “Simulation of carbon nanotube fets including hot-phonon and self-heating effects”, *IEDM Tech. Dig.*, pp. 1 4, San Francisco, CA, 10-13 Dec. 2006. [cited at p. 38]

[18] J. Appenzeller, Y.-M. Lin, J. Knoch, Z. Chen, and P. Avouris, “Comparing Carbon Nanotube Transistors—The ideal Choice: A novel Tunneling Device Design”, *IEEE Trans. on Electr. Dev.*, vol. 52, no. 12, p. 2568, 2005. [cited at p. 38, 39, 53]

[19] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris, “Band-to-Band Tunneling in Carbon Nanotube Field-Effect Transistors” *Phys. Rev. Lett.*, vol. 93, no. 19, p. 196805, 2004. [cited at p. 38]

[20] R. Grassi, S. Poli, E. Gnani, A. Gnudi, S. Reggiani,G. Baccarani, “Tight-binding and effective mass modeling of armchair carbon nanoribbon FETs”, 9-*th ULIS Conference*, p. 121–124, Udine, 2008. [cited at p. 39]

[21] G. Fiori, G. Iannaccone, G. Klimeck, "A three-dimensional simulation study of the performance of carbon nanotube field-effect transistor with doped reservoirs and realistic geometry", *IEEE Trans. on Electr. Dev.*, vol. 53, pp. 1782–1788, 2006. [cited at p. 41]

[22] G. Pennington and N. Goldsman, "Semiclassical transport and phonon scattering in semiconducting carbon nanotubes", *Phys. Rev. B*, vol. 68, no. 4, p. 045426, 2003. [cited at p. 41]

[23] H. Flietner, "The $E(k)$ Relation for a Two-Band scheme of Semiconductors and the Application to the Metal-Semiconductor Contact", *Phys. Stat. Sol. B*, vol. 54, no. 201, 1972. [cited at p. 43]

[24] J. Knoch, S. Mantl, and J. Appenzeller, "Comparison of transport properties in carbon nanotube field-effect transistors with Schottky contacts and doped source/drain contacts", *Solid-State Electronics*, vol. 49, no. 1, pp. 73-76, 2005. [cited at p. 43, 44]

[25] G. Fiori, G. Iannaccone, "Simulation of Graphene Nanoribbon Field-Effect Transistors", *IEEE Electr. Dev. Lett.*, vol. 28, pp. 760–762, August 2007. [cited at p. 44]

[26] R. Grassi, A. Gnudi, E. Gnani, S. Reggiani, G. Cinacchi, and G. Baccarani, "Hierarchical modeling of carbon nanoribbon devices for CNR-FETs engineering", *DRC* [cited at p. 45]

[27] *International technology roadmap for semiconductors 2005 edition*, available at: http://public.itrs.net. [cited at p. 48, 53]

[28] S.O. Koswatta, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, "Simulation of phonon-assisted band-to-band tunneling in carbon nanotube field-effect transistors", *Appl. Phys. Lett.*, vol. 87, p. 253107, 2005. [cited at p. 45]

[29] A. Javey, J. Guo, D. B. Farmer, Q. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, "Carbon nanotube field-effect transistors with integrated ohmic contacts and high-κ gate dielectrics", *Nano Lett.*, vol. 4, no. 3, pp. 447-450, 2004. [cited at p. 51]

[30] Y.-M. Lin, J. Appenzeller, Z. Chen, Z.-G. Chen, H. M. Cheng, and P. Avouris, "High-performance dual-gate carbon nanotube FETs with 40-nm gate length", *IEEE Electr. Dev. Lett.*, vol. 26, no. 11, pp. 823-825, 2005. [cited at p. 51]

[31] S. Koswatta, M. S. Lundstrom, and D. E. Nikonov, "Band-to-band tunneling in a carbon nanotube metal-oxide-semiconductor field-effect transistor

is dominated by phonon-assisted tunneling", *Nano Lett.*, vol. 7, no. 5, pp. 1160–1164, 2007. [cited at p. 47]

[32] J. Guo, A. Javey, D. Hongjai, and M. S. Lundstrom, "Performance analysis and design optimization of near ballistic carbon nanotube field-effect transistors", *IEDM Tech. Dig.*, pp. 703-706, San Francisco, CA, 2004. [cited at p. 51]

[33] M. Pourfath, H. Kosina, and S. Selberherr, "Geometry optimization for carbon nanotube transistor" *Solid-State Electronics*, vol. 51, pp. 1565-1571, 2007. [cited at p. 53]

[34] S.O. Koswatta, D. E. Nikonov, and M. S. Lundstrom, "Computational study of carbon nanotube p-i-n tunnel fets", *IEDM Tech. Dig.*, pp. 518–521, Washington, DC, 2005. [cited at p. 55, 59]

Chapter 3

Transport analysis in Si-NW based FETs

By means of intrinsic advantages due to their 1-D geometry, silicon nanowire structures are considered among the most interesting candidates to replace planar MOSFET devices. Reliable fabrication of such devices has been demonstrated in several experimental studies showing promising electrical performance [1, 2, 4, 3, 5, 6].

Even if theoretical studies have predicted improved electrical performance for nanowire structures due to the reduced density-of-states [7, 8] and to the non negligible contribution of ballistic electrons [9, 10, 11, 12], many other sources of scattering like thickness fluctuations [13], surface roughness (SR) at the Si–SiO_2 interface [7, 14], doping pockets and random impurities are present [15]. It has been shown that SR induces a reduction of the local density of states in the channel [7, 8], as well as an increase of threshold voltage [7], which furthermore spreads over a wide distribution for different random realizations of roughness [14, 16]. Moreover, SR-limited mobility have been calculated for SiNW-FETs using a Poisson-Schrödinger algorithm for electrostatics and then solving the Boltzmann transport equation with the Monte Carlo method [17, 18].

In order to reduce current leakage due to the small thickness of oxide layers between active channel and gate, gate stacks based on high-κ materials like HfO_2/SiO_2 or ZrO_2/SiO_2 are currently adopted for ultra-short transistors [19]. However, even if such materials provide very low gate leakage, they are probably responsible of a significant reduction of the low-field mobility. This is due to the generation of new important scattering mechanisms like remote soft phonons [20] and Remote-Coulomb Scattering (RCS) arising from the presence of additional dipoles and charged defects at the high-κ/SiO_2 interface due to the different chemical properties of two amorphous materials [21, 22, 23].

Even in the case of ultra-short nanotransistors and SiNW-FETs, low–field mobility is still considered a fundamental factor of merit. Several electrical characterizations have been performed on nanometric devices and the low-field mobility has been extracted by means of different techniques [24, 25, 26, 27, 28]. All these measurements have reported a dramatic reduction of the mobility with the channel length, whose origin is still debated. A possible explanation deals with the effect of the ballistic (or apparent) component of the mobility [29], which becomes dominant with respect to the diffusive one for short channel lengths. Such a ballistic mobility can be interpreted as generated by an ideal-contact resistance, which remains the only source of scattering in a ballistic device as formulated by the Landauer theory of transport [30, 31]. The role played by the ballistic or apparent mobility in short-channels device has been also investigated in theoretical studies [32, 33] and recently with experiments carried out on SiNW-FETs [5].

Due to the importance of tunneling current and ballistic transport in ultra-short devices a full-quantum transport simulation is envisaged to correctly describe elastic scattering mechanisms like RCS or SR scattering which are ruled by the quantum-phase coherence. In addition, another difficulty in predicting the channel mobility of such devices is that the transport of a quasi-ballistic device is strongly influenced by specific occurrences of few scattering events and hence proper averages need to be considered. In the presented study, a direct approach is used to evaluate the effect of both SR and remote-Coulomb interaction. The effective mobility is computed as the ratio between conductivity and electron density for several geometrical realizations of rough interfaces and fixed-charges distributions, evaluating the averaged impact of SR and RCS over a statistical ensemble of nanowires.

After a brief overview on the electronic properties of interest in transport simulation of silicon nanowires in Sec 3.1, the device model is presented in Sec. 3.2 with the description of the simulation parameters. The remaining sections are devoted to computational studies focusing mainly on the impact on the effective mobility of short-channel SiNW-FETs of the most relevant scattering mechanisms. In Sec. 3.3 the dependence on the device lateral dimensions of SR–limited mobility is presented. In Sec. 3.4 the impact of the remote-Coulomb scattering is investigated for different geometrical and technological parameters. Finally, in Sec. 3.5 electron-phonon interaction is considered in addition to surface-roughness and remote-Coulomb scattering and its impact on the effective mobility of scaled device is analyzed.

3.1 Electronic properties of SiNWs

Several simulation studies have been carried out on the electronic properties of silicon nanowires. Band structure calculations and applications to transport analysis have been widely reported [34, 35, 36]. Molecular dynamics simulations have been used for studies on thermal conductivity [37, 38]. Electro-mechanical and optoelectronic properties of silicon nanowire are a field of experimental analysis revealing promising applications [39, 40, 41]. As already mentioned, the effect of quantum confinement on the electron-phonon interaction as well as on the phonon-phonon interaction have been investigated [17, 42, 43]. Here, a brief overview of the major effects of quantum confinement on the electronic properties and band structure of SiNWs is given. In addition, the focus will be mainly on the effect on silicon conduction band, being the basis for the following analysis.

Bulk silicon is an indirect band gap material with a band gap of approximately 1.12 eV, with the minimum of the conduction band located in the Δ direction. Transport properties for the conduction band are mainly affected by the properties of the six degenerate Δ valleys and the information about the minima occurring in the Λ direction will be discarded. The effective mass values describing the Δ-valley ellipsoids are $\mathrm{m}_l = 0.916\,\mathrm{m}_0$ and $\mathrm{m}_t = 0.191\,\mathrm{m}_0$, where m_0 is the free electron mass.

In [35] the analysis of the major properties of the SiNWs band structure is reported. A TB approach with parameters calibrated for an accurate reproduction of both bulk silicon band gap end effective masses is used. The nanowire surface is considered passivated with hydrogen atoms, avoiding the effect of surface states associated with dangling bonds [34]. Effects of relaxation of the crystal structure are neglected using a bulk lattice constant $a_0 = 5.43$ Å [36]. The first effect to note is that the quantum confinement causes a splitting of the six-degenerate valleys of silicon. With direct reference to a square-shaped silicon nanowire with infinite transport axis along the [100] direction, four valleys are projected to the Γ point (k_x=0) of the 1-D Brillouin zone of the wire. These valleys identify the conduction band-edge and they are usually referred to as Δ_4. The remaining two valleys are zone-folded in a off-Γ position (approximately $k_x = \pm 0.37\,\pi/a_0$). They present a higher minimum energy with respect to the Δ_4 valley and are usually referred to as Δ_2 valleys or primed valleys. Concerning the valence band, a significant mixing and splitting of the band is reported [35] and it will not be considered in this discussion. In the case of an irregular cross-section, the degeneracy of the Δ_4 is removed, identifying two distinct minima doubly degenerate. This is the case in presence of surface roughness, of remote-Coulomb interaction or of a channel region partially covered by the gate, implying a separate treatment of the differently-oriented unprimed valleys.

Important consequence of the valley splitting is to transform the SiNW in a direct band gap semiconductor, with a band gap depending on the confinement itself. Both in [34] and [35] a lifting of Δ_4 and Δ_2 is reported as the size of the wire is reduced, reflected on a larger band gap and on a variation of the effective masses. The band gap is reported to be approximately 1.2 eV for a wire with a

6-nm width, increasing fast to 2.5 eV for a 1.2-nm width one.

The trend for the band gap can be understood also in terms of a wire with hard boundary conditions. Effective-mass calculations thus predict a band gap increase proportional to $\approx 1/W^2$, where W is the width of the wire. From tight-binding calculations, a deviation from this simple analytical model is found. In particular, the effective-mass prediction obtained with silicon bulk effective mass values significantly overestimate the band gap value for lateral widths less than 5 nm, as a consequence of a non-negligible increase of the Δ_4 quantization effective mass.

The study of the impact of quantum confinement on transport has been proposed in terms of transmission properties of infinite wires [35] and performance evaluation of ballistic SiNW-FET [34, 36]. It is of interest to note that in these studies there is a constant reference to both TB and EM models. In particular, the validity of the simple parabolic effective-mass approximation when applied to the analysis of SiNW-FET is investigated. Directly connected to the previous observations on the dependence of the band gap on the wire width, it is clearly found that the EM approach, calibrated with bulk parameters, gives an incorrect transport description as the wire dimensions are scaled down to few nanometers (below 5 nm). This implies a redefinition of the transverse and longitudinal mass values in order to successfully extend the application of this method. With the use of properly calibrated values the tight-binding $I - V$ curves can be nicely reproduced over a wide range of wire widths [36]. In the following, values for the transverse and longitudinal effective mass different from the bulk ones will be used for $W < 5$ nm. Other effects on the band structure can not be in any case captured by the effective-mass approximation. In particular, the resolution of the valley degeneracy for both Δ_4 and Δ_2 valley as the wire dimensions are shrunk down to 2 nm and below. Narrower nanowires show also a larger dependence of the valley degeneracy on the gate voltage. In addition, channel orientations different from the [100] analyzed so far, are less accurately described by the effective-mass approach [44]. The presented results in any case confirmed the validity of the parabolic effective-mass approach for the cases presented in the following sections, as it captures all the essential transport features.

3.2 Device Model

Two two-dimensional sketches of the device are reported in Fig. 3.1: one (top) representing a device cross-section along the transport direction; the other (bottom) a cross-section on transverse yz-plane. The device is supposed to be infinite along the transport direction, through the inclusion, with the model previously described, of ideally ohmic contacts extending the doped source and drain regions reported in figure. A uniform doping distribution with abrupt junction with the

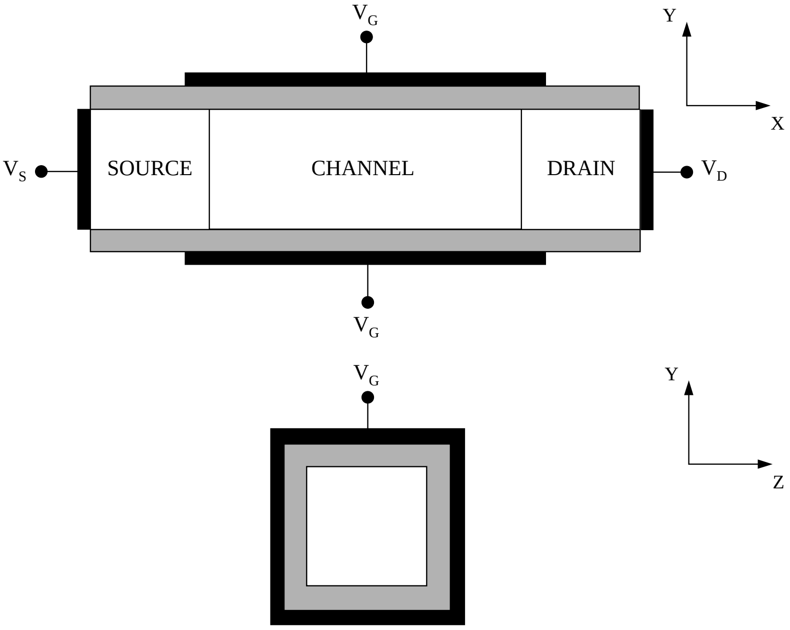

Figure 3.1: *(Top) Cross-section of the device along the transport direction, x. The subdivision of the device in three regions, source, drain, and channel, is highlighted. (Bottom) Cross-section on the transverse yz-plane. A GAA architecture is supposed.*

intrinsic channel is supposed at the contacts. Only the case of n-type devices is considered due to known limitations of the effective-mass model in describing transport in silicon valence band. An oxide layer covers the entire nanowire surface.

Several experimental realizations have been reported during the last years for SiNW-FETs, showing different architectural choices for the gate electrode. Solutions moving from partially covered channels, as π-gate or Ω-gate FETs, to GAA devices, have been reported. Although the possibility of activating independently the four gate electrodes have been considered in the numerical solver, all simulations will refer to GAA devices, known to ensure the best electrostatic control reducing short channel effects.

All the results that will be presented, in both ballistic and dissipative condition, have been obtained by solving the transport problem in the Coupled Mode Space approach. The eigenfunctions to be used as basis set for the expansion of the real-space Green's functions are calculated for each slice of the device normal to the transport direction, by imposing vanishing boundary conditions in the outer mesh nodes, including the wave function penetration in the oxide. In presence of surface roughness or remote fixed charges, up to 50 transverse modes have been considered to have a good convergence to the real-space solution. The saving in computational time is still in any case high, allowing for the simulation of devices with a total cross-section extension (silicon+oxide) up to 11 × 11 nm^2 and total length up to 60 nm. In order to optimize the solution of the sparse eigenvalue problem in the transverse plane, the routines of the ARPACK package have been used [45].

The solution of the Schrödinger equation is performed within the parabolic effective-mass approximation. Focusing on the linear regime analysis of transport (low V_{DS} biases) the parabolic effective-mass approximation is still expected to give realistic results, although quantitative variations on the effective mobility value have been reported using a non parabolic model [46]. The channel orientation is supposed along the [100] direction. Bulk silicon effective masses (m_l = 0.916 m_0, m_t = 0.191 m_0) are used for wires with a lateral section wider then 4 nm, while corrected masses (m_l = 0.960 m_0, m_t = 0.230 m_0) are adopted for the smaller sections [34]. When describing penetration into the gate dielectric, the oxide effective mass is fixed at one half of m_0. The fundamental equations to be solved in the transport analysis are the same presented in Sec. 1.3 for the electron transport in conduction band within the CMS approach.

The electrostatic potential is calculated by solving the 3-D Poisson equation. In accordance with the constraint imposed by the treatment of the surface roughness (see Sec. 3.3), a uniform discretization step of 0.2 nm is used for the three spatial directions. The total charge density used for the solution of the Poisson equation accounts for electrons, source and drain doping impurities, and for the oxide fixed charges, when included. As the analyzed devices have an intrinsic channel and n-doped contacts, the hole contribution to the total charge is discarded. The contribution of the exchange-correlation potential to the self-consistent electrostatic potential is included as well. The additional exchange-correlation potential, V_{xc}, due to the electron-electron interaction is considered following [47] as:

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{q}{8\pi\epsilon_{si}a}\frac{2}{\pi\alpha r_s}\left[1 + 0.7734x\ln\left(1+\frac{1}{x}\right)\right], \tag{3.1}$$

where $r_s = a^{-1}[4\pi n(\mathbf{r})/3]^{-1/3}$, $x = r_s/21$, $\alpha = (4/9)^{1/3}$, $a = 4\pi\epsilon_{si}\hbar^2/(m^\star q^2)$ is the effective Bohr radius, and $n(\mathbf{r})$ is the local electron density. As proposed in [47] and works cited therein, the aim of such potential correction is a qualitative inclusion of this many-body effect.

The parameters for the calculation of the electron-phonon energy are presented in Table 3.1. In addition to acoustic phonons, a g-type and two f-type optical modes are considered, as they are the most dominant contributions to transport. Parameters are taken from [48], apart for a slight modification in the phonon energy for the two f-type modes, as already described in [49]. The modification is due to numerical considerations, allowing for a reduction of the computational costs. The inclusion of the electron-optical phonon interaction, as mentioned above, implies a coupling between the solution at energies at a distance

Ξ	$\mathrm{D}_t\mathrm{K}_j$ (10^8 eV/cm)	$\hbar\omega_j$ (meV)	selection rule
14.6			intravalley
	11.0	63.2	g-type
	2.0	47.4	f-type
	2.0	63.2	f-type

Table 3.1: *Parameters for the electron-phonon coupling.*

$\Delta E = \hbar\omega_0$, where ω_0 is the frequency associated to the specific optical mode. In order to include every optical-phonon coupling, a sufficiently refined energy grid is required, along with the simultaneous storage of the coupled Green's functions during the iterative solution of the kinetic equations in the self-consistent Born approximation. For the chosen parameters, a fixed minimum energy step of $\Delta E_{\min} = 15.82$ meV is identified, being the phonon energies of the considered modes all multiple integers of $\Delta E_{\min}$. By this way, it is possible to describe the coupled solution for energies ranging from a minimum value $E_{\min}$ to a maximum value $E_{\max} = E_{\min} + N \times \Delta E_{\min}$, with N integer, accounting for the complete set of relevant phonon modes. Additional refinements to the solution are obtained by considering multiple nested energy grids having a lower energy value in the range $[E_{\min}\ , E_{\min} + \Delta E_{\min}\,[$. For the analyzed cases, reasonable values of N are found in a range of maximum 40 giving accurate solutions and allowing for the simulation of devices with large dimensions. An optimization of the number of sub-grids to be considered is obtained through the use of Gaussian quadratures. Typically, a set of 20 sub-grids is used.

3.3 Size dependence of SR–limited mobility

As previously remarked, surface roughness is often referred to be the most important cause of mobility degradation in conventional MOSFETs at high transverse fields. In SiNWs several competing mechanisms determine the impact of SR. As reported in [7], a reduction of the density-of-states, in association with the effect of volume inversion [17], will lead to a beneficial reduction of SR impact on mobility. On the other side, as the dimension of the wires is reduced, in connection to enhanced potential fluctuations, the SR scattering becomes stronger at low transverse fields, causing a drastic mobility decrease [46].

In the following, surface roughness is not treated by using a perturbative method, as widely described in the literature with reference to the solution of the Boltzmann equation. A microscopic treatment of potential fluctuations will be considered, assuming a specific realization of rough interfaces between the silicon and silicon-oxide in accordance with a given statistical model. The model has already been used in [7] and [14] for the case of full-quantum simulations within

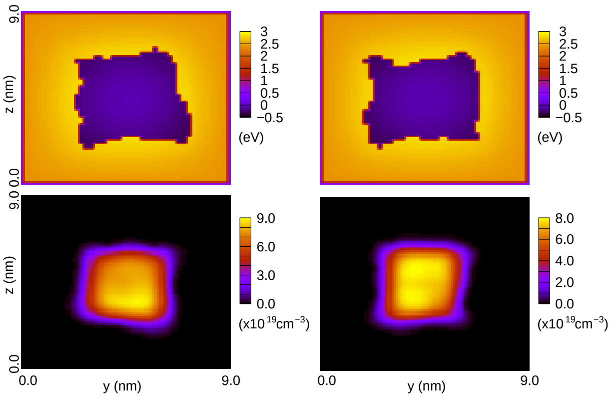

Figure 3.2: *(Top) Conduction band profiles for two sections of the channel (bottom) and the corresponding electron densities. The wire width is* 5 *nm. SR parameters are:* $\Delta_m = 0.3$ *nm and* $L_m = 1.0$ *nm.*

the parabolic effective-mass and tight-binding approximation, respectively. In the presented work, the importance of a statistical approach in the evaluation of device performance is emphasized, focusing on the effect of such fluctuations on the effective mobility in devices with different lateral dimensions.

3.3.1 SR generation

An abrupt, randomly varying interface between Si and SiO_2 is considered, using a two-parameters autocorrelation model. According to [50] a quasi-continuous function $\Delta(\vec{\mathbf{r}})$ representing the two-dimensional Si–SiO_2 interface displacement is assumed to be statistically characterized by an exponential autocovariance function

$$C(\vec{\mathbf{r}}) = \Delta(\vec{\mathbf{r}}\,')\Delta(\vec{\mathbf{r}}\,' - \vec{\mathbf{r}}) = \Delta_m^2 e^{-\sqrt{2}\mathrm{r}/L_m} \ , \tag{3.2}$$

where Δ_m is the Root Mean Square (RMS) of the fluctuations and L_m the correlation length. The accuracy of the above model has been validated in [50] by means of an experimental characterization of planar Si(100)-SiO_2 interfaces. Each particular realization of SR was generated by starting from the Power Density Spectrum (PSD) obtained by transforming the correlation function (3.2) and adding a random odd phase to the square root of the power spectrum in order to create a statistics of local fluctuations in the momentum space. The real-space realization of the rough surface is obtained by anti-transformation and then by a uniform discretization with a 0.2 nm step. In order to ensure a better quality of the statistical properties of the interfaces, four surfaces bigger than the ones needed for the definition of the channel of the device are generated and subsequently cut to the actual device dimensions and merged.

An example of the interfaces resulting from this random generation is depicted in Fig. 3.2. Contour plots (top) of the conduction band profile are reported for

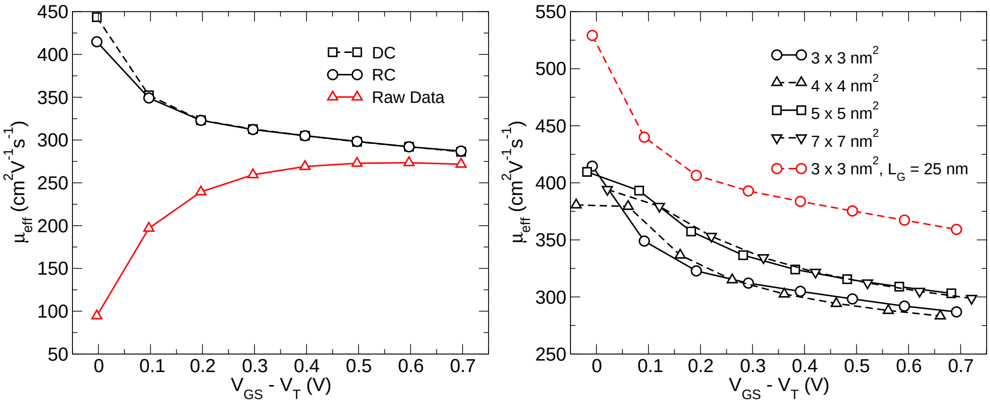

Figure 3.3: *(Left) Effective mobility data correction for a device with ideal Si–SiO$_2$ interfaces. Raw data: whole gated region extension integration. DC: differential charge evaluation. RC: reduced channel region integration. In the RC approach 5 nm at the left and right contact-to-channel interfaces are removed in the density evaluation. The wire section is 3×3 nm^2 and $L_G = 20$ nm. (Right) Ballistic mobility for different wire sections and fixed channel length, $L_G = 20$ nm (black lines, filled and open symbols) and for a 3×3 nm^2 wire section with $L_G = 25$ nm (red line with symbols).*

two different slices along the transport direction. The device section is of 5×5 nm^2 and the roughness parameters are set to $\Delta_m = 0.3$ nm and $L_m = 1.0$ nm. Surface roughness affects differently the two slices by creating variations in the confining potential well. Direct consequence is the change in shape of the eigenfunctions given by the solution of the 2D Schrödinger equation on the transverse plane. The change of the eigenfunctions causes the rise of the coupling effect between transverse modes wich heavily impacts the electron properties of the device. The resulting electron densities in the two slices are reported as well (bottom), showing a different deformation of the charge distribution.

In order to focus on the impact of roughness on the mobility of SiNWs, rough SiSiO$_2$ interfaces are generated only in the gated region. Interfaces are kept ideal in the source and drain regions. No coupling between different valleys has been considered as the surface-roughness scattering is an intra-valley scattering mechanism [46].

3.3.2 Effective mobility extraction

Effective mobility extraction has been performed in a linear transport regime. Therefore the effective mobility reads:

$$\mu_{\text{eff}} = \frac{G L_{\text{ch}}}{q N_{1D}} , \tag{3.3}$$

where G is the conductance, L_{ch} the channel length and N_{1D} the channel electron density per unit length.

A crucial theoretical point is the correct definition of the charge density in (3.3), that is assumed to be uniform in the whole channel region. A calculation of the density over the entire gated region gives rise to spurious effects due to carrier penetration from the contact regions (gate-to-contact capacitances). This is shown in Fig. 3.3 (left) as Raw Data in the case of an ideal ballistic device with a cross section of 3×3 nm^2 and gate length $L_G = 20$ nm. For all the reported simulations, a low drain bias of $V_{\mathrm{DS}} = 5$ mV is assumed and the temperature is fixed at T = 300 K. Such a charge injection constitutes an offset overwhelming the contribution of the inversion charge as V_{GS} decreases, causing a drastic reduction of the extracted mobility values. In order to eliminate the gate-to-contact charge component two different methods have been proposed and compared.

First, the use of a differential charge calculation was considered [26]. Two close gate lengths are considered ($L_G = 20$ and $L_G = 25$ in this study) and a differential electron density as a function of the channel length is calculated as

$$N_{1D}^{(0)}(L) = \frac{N_{1D}(L')\,L' - N_{1D}(L)\,L}{L' - L}\ , \tag{3.4}$$

where L', and L are the two channel lengths.

As long as shifts in the threshold voltage (V_{T}) are negligible, the same electrostatic behavior and hence the same gate overdrive for a given gate bias, V_{GS}, are supposed for the two devices. Using the density given by (3.4) in the calculation of (3.3) the corrected mobility curve labeled “differential charge” (DC) in Fig. 3.3 is obtained. In such a way the ballistic mobility curve recovers a monotonic decrease as a function of the gate bias as expected, which is due to the effect of the Fermi statistics.

The second approach is based on the reduction of the integration region to be considered for the calculation of the electron density. In this way, the density evaluation is restricted over a section of the gated region characterized by an almost uniform electron distribution not affected by the diffusion of carriers at the contact-to-channel interfaces. The result obtained with this method is reported in Fig. 3.3 labeled as “reduced channel” (RC). An overall good agreement with the differential approach is found. Differences of small magnitude are present only in the sub-threshold regime where the mobility is extremely sensitive to the charge values. Anyway the results are unaltered from the qualitative point of view.

The same analysis has been conducted taking into consideration devices with rough Si–SiO$_2$ interfaces. In this case a further difficulty comes from the threshold voltage fluctuation for different surface roughness realizations [14]. This implies that in the sub-threshold regime the charge does not always linearly depend on the channel length, which makes the DC extraction impractical. This effect is

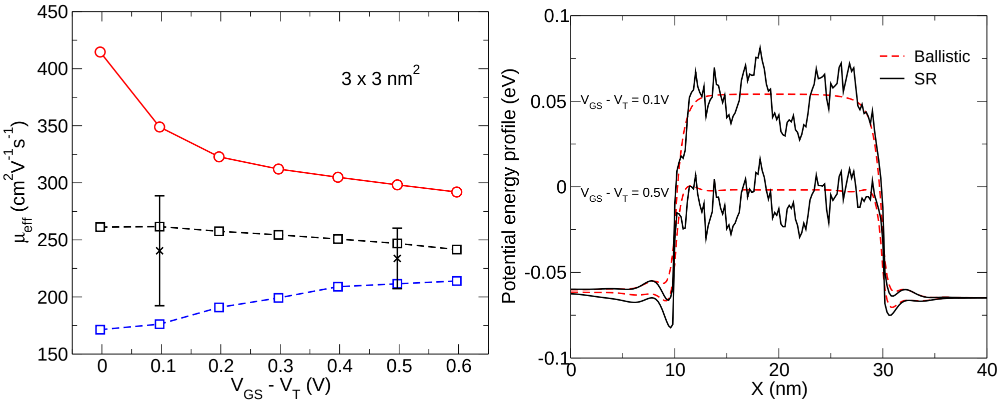

Figure 3.4: *(Left) Ballistic mobility (solid red line with symbols) and SR–limited mobility (dashed black and blue lines with symbols) for a* 3×3 *nm*^{2} *section wire. Expectation value and sample standard deviation are reported (dispersion bars). The evaluation has been conducted over* 20 *different geometrical realizations. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm. (Right) Potential energy profiles for the first subband of an unprimed valley for two different bias conditions (*$V_{\rm GS} - V_T = 0.1$ *V and* $V_{\rm GS} - V_T = 0.5$ *V) with ideal (dashed red line) and rough (solid black line) interfaces. Device and SR parameters as above.*

particular strong for smaller wire sections, while the expected behavior is recovered when moving to larger sections and higher gate voltages. Therefore, in the following the SR–limited mobility data have been extracted with the RC integration method.

3.3.3 Ballistic mobility analysis

Preliminary simulations have been performed to extract the ballistic or apparent mobility in devices with different lateral sizes. The ballistic mobility has been used both as a reference case and as a first validation of the used methodology. Then, the effect of the SR on the effective mobility has been calculated and analyzed by a comparison with the ideal ballistic case.

The analyzed devices have square sections with channel width ranging from 3 to 7 nm. The gate length is fixed at 20 nm. The silicon-oxide is 2 nm thick with ideal Si–SiO$_2$ interfaces. Source and drain contacts with a donor doping concentration of 2×10^{20} cm^{-3} and a conduction band-edge gate work-function are considered.

The obtained results are depicted in Fig. 3.3, where the mobilities of SiNWs with four different sections are reported as a function of the gate overdrive, $V_{\rm GS} - V_T$. First, a monotonic decrease of mobility is noted as gate voltage increases. This effect can be ascribed to the increasing importance of the electron

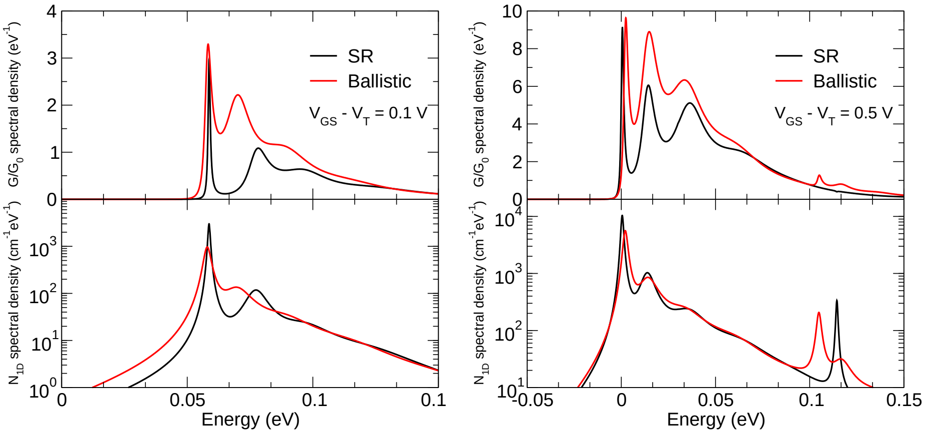

Figure 3.5: *(Top panel) Conductance spectral density and (bottom panel) channel charge density evaluated for the potential profiles reported in Fig. 3.4 in both (red lines) ballistic and (black lines) SR case. The gate biases are (left)* $V_{\mathrm{GS}} - V_T = 0.1$ *V and (right)* $V_{\mathrm{GS}} - V_T = 0.5$ *V.* G_0 *is the quantum conductance, accounting for valley and spin degeneration. The wire section is* 3×3 *nm*^{2}*. The drain bias is fixed at* $V_{\mathrm{DS}} = 5$ *mV. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm.*

gas degeneracy. In fact, as expected from an analytical flux theory for a single-occupied 1-D subband [10],

$$\mu_{\mathrm{eff}} \propto \frac{\mathcal{F}_{-1}(\eta_F)}{\mathcal{F}_{-\frac{1}{2}}(\eta_F)} \,, \tag{3.5}$$

where $\mathcal{F}_n(\eta)$ is the n-th order Fermi integral and $\eta_F = (\mu_s - E_0)/(kT)$, with E_0 being the subband eigenvalue.

As the primed valley begins to be populated, a degradation of mobility is also caused by its larger effective mass in the transport direction. This effect is evident when comparing the mobility values of the 3×3 with the 4×4 nm^2 wire and of the 5×5 with the 7×7 nm^2 wire for large gate voltages. The jump in the values when passing from the 4×4 nm^2 to the 5×5 nm^2 wire is due the change in the quantization condition (effective mass variation).

As the wire width is additionally increased, an almost equivalent contribution to mobility is obtained from the populations of the primed and unprimed valleys and the mobility saturates to a section-independent value. In addition, the ballistic mobility for a 3×3 nm^2 section wire with L$_G$ = 25 nm is reported. As expected, a perfect linear dependence on the gate length is found.

3.3.4 Impact of surface roughness

The study of the impact of surface roughness on the device performance has been carried out through a comparison of the SR–limited with the ideal ballistic mobility. Three different cross-sections are deeply analyzed, with 3, 5 and 7 nm lateral width. The gate length is fixed to $L_G = 20$ nm, focusing, in this way, on the quasi-ballistic transport regime. SR parameters are set to $\Delta_m = 0.2$ nm and $L_m = 1.0$ nm, which are expected to be typical values for Si–SiO$_2$ interfaces [50]. The other electrical and physical parameters are the same as above.

The analysis has been conducted following a statistical approach, taking into consideration different surface realizations in order to estimate the expectation values and their standard deviation. This has been necessary since the considered gate length was not long enough to reach a complete diffusive transport regime.

In Fig. 3.4 (left) the results obtained for a 3×3 nm^2 wire are reported. The solid red line refers to the ideal ballistic case while the dashed are two different SR–limited mobility curves reported as examples. The expectation value and the sample standard deviation reported at low and high gate bias conditions have been evaluated over 20 different geometrical realizations. The values are reported as a function of $V_{\mathrm{GS}} - V_T$, where V_T is the ballistic threshold voltage, not accounting for the possible shift in V_T in each single realization.

For the 3×3 nm^2 wire case, although a reduction of mobility is obtained in presence of SR with respect to the ballistic case, a non-uniform trend of the mobility curves as a function of the gate voltage is found, depending on the specific surface realization. This non trivial dependence is exhibited by the two examples in Fig. 3.4, one showing a slight mobility reduction, the other a mobility increase from sub-threshold to high biases regime.

Such behaviors imply a larger mobility variation for low rather than for high voltages, evidencing a reduction of the importance of SR with the increase of the channel carrier density. This result is opposed to the classical interpretation of the SR effect and can be ascribed to the effect of the quantum-phase coherence, that plays an important role when energy independent scattering mechanisms are considered.

Indeed, one of the two main effects of surface roughness is to create large fluctuations of the effective potential profile experienced by the electrons, the other being the enhanced coupling strength between transverse modes of adjacent sections [7]. In Fig. 3.4 the potential profile of the first subband of an unprimed valley is reported at low and high gate biases both for the SR and the ballistic case, showing that the potential fluctuations are the predominant effect for the 3×3 nm^2 section wire.

Several potential wells can be identified along the transport direction, suggesting the occurrence of charge localization. In order to investigate the latter phenomenon, the two physical quantities used in the calculation of μ_{eff}, that are

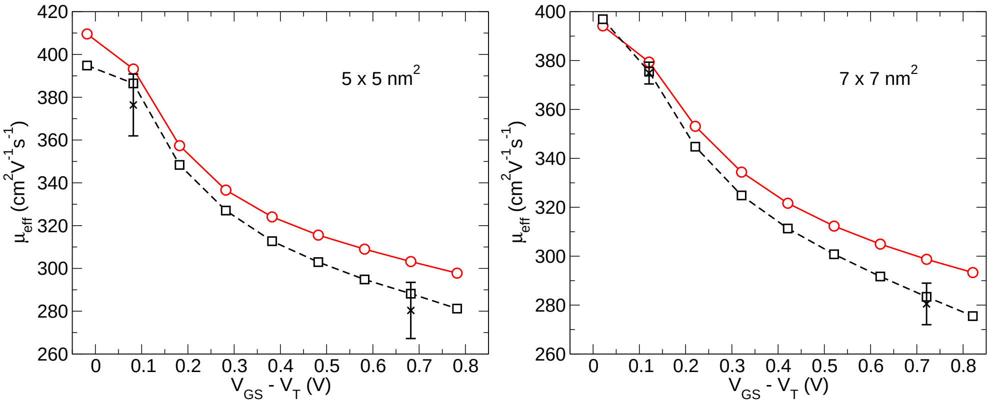

Figure 3.6: *Ballistic mobility (red line with symbols) and SR–limited mobility (dashed black lines with symbols) for a (left)* 5×5 *nm*^{2} *and a (right)* 7×7 *nm*^{2} *section wire. Expectation value and sample standard deviation are reported (dispersion bars). The evaluation has been conducted over* 20 *different geometrical realizations. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm.*

the conductance and the carrier spectral densities, have been analyzed. In the top panel of Fig. 3.5 they are shown for $V_{\mathrm{GS}} - V_T = 0.1$ V, comparing the ballistic and the SR case. The same parameter set used in the calculation of Fig. 3.4 (right) was adopted. SR strongly modifies the shape of the two spectral densities characterized by sharper peaks with respect to the ballistic case. At low gate voltages, this gives rise to a detrimental reduction of the conductance, whereas charge density can even be increased. Hence, the SR–limited mobility value is found to be reduced down to the 50% of the ballistic value.

As the gate voltage is increased and further propagating states appear, localization effects become less important and the fluctuations of the potential profile are self-consistently smoothed, resulting in the widening of the wells. As shown in the bottom panel Fig. 3.5 the differences between the spectral densities in the ballistic and SR case become less important. Depending on the specific realization of SR this can result in a possible increase of mobility with the gate voltage. In any case, the trend at high gate voltages is characterized by mobility values less spread than at low V_{GS}, with a reduction of the standard deviation with the increase of the gate voltage that is found to scale from 20% to 11.3% of the expectation value.

A similar analysis has been carried out for wires with lateral widths of 5 and 7 nm. The results are shown in Fig. 3.6 where a specific SR–limited mobility curve is reported compared to the ideal ballistic case for the two different geometries. In contrast with the previous result for the 3×3 nm^2 section wire, a monotonically decreasing mobility for increasing gate voltages is reported in both cases.

Due to the widening of the wire section for a fixed RMS value in the surface

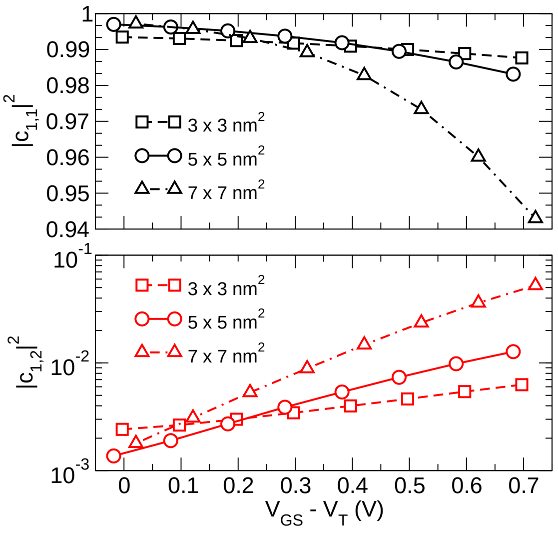

Figure 3.7: *Mean value of the squared modulus of the autocorrelation of (top) the first mode of an unprimed valley and of (bottom) its correlation with the second mode evaluated along the channel for a given roughness realization and different device sections (different symbols) as function of the gate voltage. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm.*

generation, the effect of the quantum confinement is reduced with respect to the previous case and smoother subband profiles are obtained (not shown). Reduced potential fluctuations cause a smaller degradation of the conductance reflected by an overall higher ballisticity and by the suppression of the potential wells in the channel. This means that transport properties of such devices are mainly affected by the scattering mechanism through the transverse mode coupling. This effect becomes more important as the gate overdrive increases and turns on additional states which contribute to the channel density and to the mode coupling, leading to the mobility degradation.

In Fig. 3.7, in order to estimate the impact of mode coupling, the mean value along the channel of the autocorrelation of the first mode of an unprimed valley ($|\mathrm{c}_{1,1}|^2$) and of its correlation with the second mode ($|\mathrm{c}_{1,2}|^2$), as defined by (1.22), have been computed. As expected, the correlation coefficients $|\mathrm{c}_{1,1}|^2$ and $|\mathrm{c}_{1,2}|^2$ show a stronger dependence on V_{GS} for SiNWs with larger sections.

A further contribution to this analysis is presented in Fig. 3.8, where the charge densities for the three considered sections are displayed at the same large gate overdrive. Here, a drastically different distribution of the charge is shown, with the charge in the smallest section confined in the central region, whereas it is pushed at the surface for the largest one. This is also consistent with the expected classical behavior predicting that the increasing importance of SR limited mobility for large gate voltages is due to the charge squeezing toward the Si–SiO$_2$

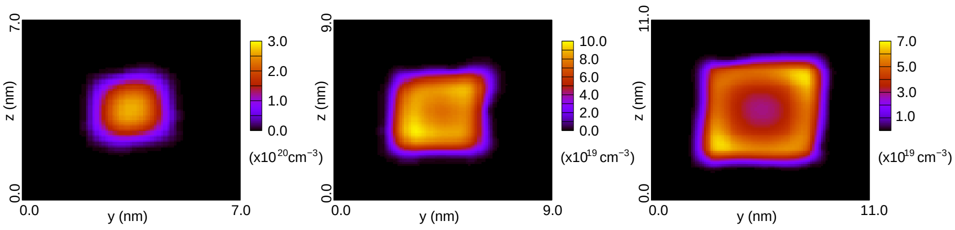

Figure 3.8: *Charge density computed for a given slice of the channel and different lateral widths (3, 5, 7 nm from left to right). The densities are evaluated at the same* $V_{\mathrm{GS}} - V_T \approx 0.7$ *V condition. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm.*

interface.

It is also of interest to note that, due to the increasing importance of the coupling effect over that of the potential profile deformation, at larger sections a reduced standard deviation of the mobility values is obtained (averaging effect). At high gate bias it reduces from 11.3% of the mean value for the 3×3 nm^2 wire to 4.6% and 3.3% for the 5×5 and 7×7 nm^2 ones, respectively.

By comparing the two limiting scattering mechanisms, one finds out that potential fluctuations are almost insensitive to the gate voltage, emerging in particular at low gate biases, whereas the mode-mixing strongly depends on the bias voltage and becomes more and more important as the gate voltage is increased.

Finally, the previous data have been collected to carry out a statistical analysis of the ratio between the SR–limited mobility and the ideal ballistic mobility at low and high gate biases. In Fig. 3.9 the probability distribution of such mobility ratio over 20 different geometrical realizations of the interfaces is reported. This figure summarizes the main results of this analysis, showing the general reduction of the mobility dispersion as gate voltage is increased and the larger influence of SR for smaller wires.

It is important to notice that the presented results quantitatively depend on the choice of the RMS of the surface roughness. If a larger value of RMS is used, the impact of the potential-fluctuation mechanism would become more and more predominant, especially at low gate biases. This is somehow confirmed by the analogy with ultra-thin body SOI transistors, where such effect has been experimentally observed [13].

3.4 Impact of remote-Coulomb scattering

The RCS–limited mobility has been theoretically investigated starting from scattering rates modelled within the Born approximation in [51] and electron mobility degradation due to RCS has been calculated in ultra-thin oxide MOSFETs

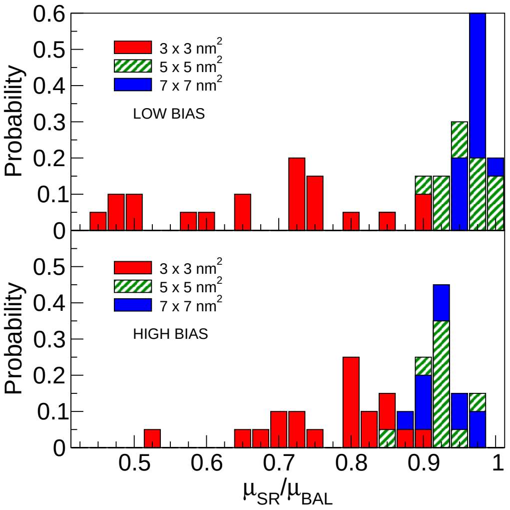

Figure 3.9: *SR–limited over ballistic mobility probabilities at (top) low and (bottom) high gate biases for the three different geometries analyzed. Statistics collected over* 20 *different roughness realizations. SR parameters are:* $\Delta_m = 0.2$ *nm and* $L_m = 1.0$ *nm.*

by Esseni and Abramo in [52] by using the relaxation-time approximation and showing the important role of the electron screening. Saito *et al.* proposed an analytical model reproducing the mobility reduction due to fixed trapped charges at the Al_2O_3/SiO_2 interface of metal-insulator-semiconductor field-effect transistors (MISFETs) and the mobility increase with the interfacial layer thickness t_{IL} [1, 53]. Numerical calculations based on the Monte-Carlo method have been carried out to compute the RCS–limited mobility and its reduction as measured on high-κ/SiO_2 gate stack MOSFETs [54, 55].

Due to the use of ideal metallic gates, the effect of the remote-Coulomb interaction will be considered only in presence of a high-κ/SiO_2 gate oxide stack, with fixed-charge centers localized at the interfacial plane between the two different dielectrics. As for the case of surface roughness, the approach is non-perturbative, considering different distributions of impurities and statistically evaluating the impact on the effective mobility.

3.4.1 Fixed charges generation

In Fig. 3.10 a schematic section of the device in the gated region is reported. An interfacial layer of silicon oxide with thickness t_{IL} is placed between the silicon channel and the high-κ gate dielectric. Positive and negative impurities are supposed to be randomly localized at the interface with a uniform probability distribution over the 2-D interfacial plane with a given density N_{Fix}.

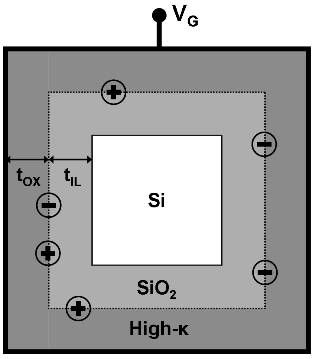

Figure 3.10: *Sketch of the transverse section of the device under the gate region. A* SiO_2 *interfacial layer, with thickness* t_{IL}*, lays between the high-*κ *oxide and the silicon channel. Randomly distributed positive and negative charges are present over the interfacial plane.*

The presence of fixed charges is considered only in the region covered by gate, thus allowing to focus on channel transport properties and to study the impact on effective mobility. No coupling between modes arising from different valleys is considered, as the remote-Coulomb interaction is an intra-valley scattering mechanism [52].

This model has been already proposed in [56] and applied to the study of planar thin body transistors. In [56] the use of fixed charges of opposite sign is at the basis of the observed mobility degradation connected to remote-Coulomb scattering. A uniform charge distribution located at the gate/oxide interface is indeed supposed to generate a simple shift of flat-band voltage with small impact on the channel mobility due to the distant location. On the other side, dipoles located at the interfacial layer are able to affect the carrier transport through potential oscillations, while no significant shift of the flat-band voltage is expected due to a quantitative balance between charges of opposite sign. The magnitude and impact of the potential oscillations strongly depend on the choice of the specific values of N_{Fix} and t_{IL}, being the only two parameters introduced in the model.

It is important to recall that in the presented model a statistically uniform but discrete distribution of charges is introduced. Thus a dependence on the specific realization of the charge distribution is expected on the transport properties of the device. The evaluation of the RCS effect on the effective mobility will be carried out by following a statistical approach and averaging the results obtained on a set of realizations.

In the following, rectangular shaped Si-NWs with a fixed gate length of 40 nm and fixed lateral dimensions of 5×5 nm^2 will be considered. Source and drain regions are supposed 10 nm long and doped with a donor concentration of 2×10^{20}

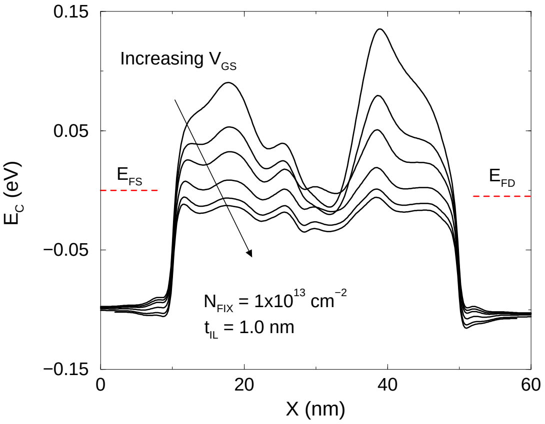

Figure 3.11: *Energy profiles along the transport direction for the first mode of one of the unprimed valleys for different gate bias conditions. RCS parameters are:* $N_{\mathrm{Fix}} = 1 \times 10^{13}$ *cm*$^{-2}$ *and* $t_{\mathrm{IL}} = 1.0$ *nm. The source and drain Fermi levels,* E_{FS} *and* E_{FD}*, are reported.*

cm^{-3}. A conduction band-edge gate work-function is used. As in the previous analysis, the drain voltage is fixed at $V_{\mathrm{DS}} = 5$ mV and a gate voltage sweep from the sub-threshold regime to the high overdrive state will be considered.

The high-κ oxide thickness t_{ox} is 2 nm, while t_{IL} varies from 0.6 to 1.4 nm. The dielectric constant of the external layer has been set to twice the value of the silicon-oxide. The chosen value of t_{ox} allows for neglecting the effects of the gate leakage current and ensures a good convergence of the deformed potential profiles to the boundary conditions fixed by the ideal metallic gate. The density of the fixed charge centers on each of the 2-D interfacial planes, N_{Fix}, is varied in the range from 5×10^{12} cm^{-2} to 2×10^{13} cm^{-2}. The ballistic device, used as reference, is characterized by the absence of trapped impurities.

3.4.2 Screening effect

The main effect of the presence of fixed-charge centers on the transport properties of the device is to induce spatial fluctuations of the potential profile and hence of the 1-D subbands experienced by the electrons. The related variation of the confinement condition at each slice of the device along the transport direction causes an additional enhancement of the coupling effect between transverse modes. The two competing mechanisms affect the carrier transport with different strength at different gate biases.

The inclusion of the impurities directly in the Poisson equation as additional charge-density contribution gives a solution that self-consistently accounts for the screening effect due to the channel carriers. This is clearly shown in Fig. 3.11 where the energy profiles of the first mode of one of the unprimed valleys are

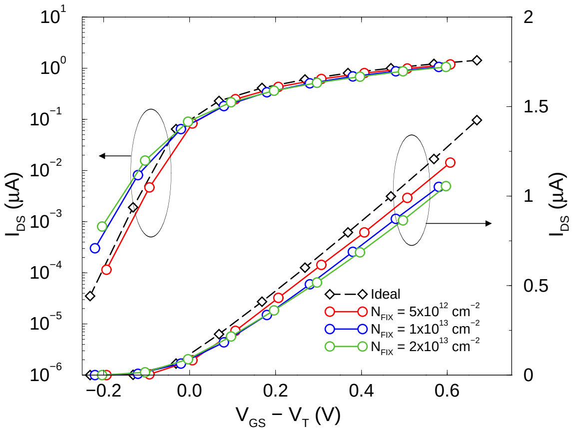

Figure 3.12: *(Solid lines with symbols) Turn-on characteristics for devices with different* N_{Fix} *values and a fixed value of* $t_{\mathrm{IL}} = 1.0$ *nm. (Black dashed line with symbols) The reference ballistic case is reported for comparison. The drain voltage is set to* $V_{\mathrm{DS}} = 5$ *mV.*

reported for different gate bias conditions. As the gate voltage is increased and a higher carrier density is present in the channel, the Coulomb potential generated by the fixed charges is screened by the electrons at the surface, thus reducing the perturbation effect. A progressive smoothing of the potential energy fluctuations is shown, while the potential profile gets closer to the ballistic flat one.

As previously stated, at low gate voltages, the presence of marked and sharp oscillations of the potential is the limiting scattering mechanism. At higher gate overdrives the confinement of the carriers near the Si/SiO$_2$ interfaces induces an increased mode mixing, which prevents the transport solution to recover a purely ballistic condition. This effect and its dependence on the chosen lateral size of the wire has been already investigated in the previous section in presence of surface roughness. It is important to note that in the case of rough Si/SiO$_2$ interfaces no significant role of the screening effect on the subband deformation was reported.

3.4.3 Transfer characteristics analysis

A preliminary analysis is focused on transfer the characteristics at low drain bias and different values of the interface layer thickness and impurity density. A qualitative description of the impact of the remote-Coulomb interaction on the I_{OFF} current and on the threshold voltage is also reported.

In Fig. 3.12, the output characteristics in both linear and logarithmic scale are reported for devices with $N_{\mathrm{Fix}} = 5\times10^{12}$, 1×10^{13}, and 2×10^{13} cm^{-2}, considering a single realization of impurity charge distribution for each parameter value. The interfacial layer thickness is kept at $t_{\mathrm{IL}} = 1.0$ nm. The proposed values for the fixed charges density are chosen as typical values presented in the literature and

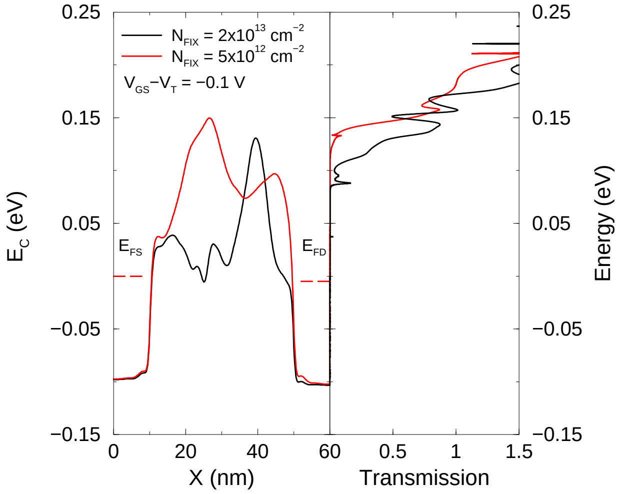

Figure 3.13: *(Left) Potential energy profiles of the lowest unprimed subband for two different values of N_{Fix}, namely (red) 5×10^{12} and (black) 2×10^{13} cm^{-2}, at approximately the same gate overdrive $V_{\mathrm{GS}} - V_T \approx -0.1$ V. (Right) The transmission probabilities are reported for this configuration (horizontal axis).*

are consistent with flat-band shift measurements [1]. The curves are plotted as a function of $V_{\mathrm{GS}} - V_T$, where V_T is the threshold voltage extracted from the linearization of the output characteristic at high overdrive. First, a progressive increase of V_T with increasing defect densities is observed. As it is stated in the description of the model for the inclusion of charged centers, the random generation of discrete fixed charges deviates from the ideal condition of a uniform distribution of dipoles causing a non-compensation of the impact on the potential itself. This is especially observed at higher densities which lead to the formation, with high probability, of positive/negative charge clusters. An example of such a phenomenon is reported in the left panel of Fig. 3.13, where the subband profiles at a sub-threshold bias for devices with $N_{\mathrm{Fix}} = 5 \times 10^{12}$ cm^{-2} and $N_{\mathrm{Fix}} = 2 \times 10^{13}$ cm^{-2} are shown. An increase of the threshold voltage of 60 mV is found for the device with $N_{\mathrm{Fix}} = 5 \times 10^{12}$, whereas $\Delta V_T = 170$ mV for the device with $N_{\mathrm{Fix}} = 2 \times 10^{13}$ cm^{-2}. Such anomalous increase reflects the effects of a cluster of positive charges in the device with $N_{\mathrm{Fix}} = 2 \times 10^{13}$ cm^{-2}.

Moreover, the presence of charge clusters has a direct effect on the current characteristics. In fact, depending on the specific realization of fixed charges, favorable tunneling paths can be possible. The tunneling component of the current has a detrimental impact on the performance of the device with a sharp increase of the current levels in the off-condition and with a reduction on the gate control capability, reflected by a progressive increase of the inverse sub-threshold slope.

By decreasing the gate voltage and depleting the channel, higher barriers are

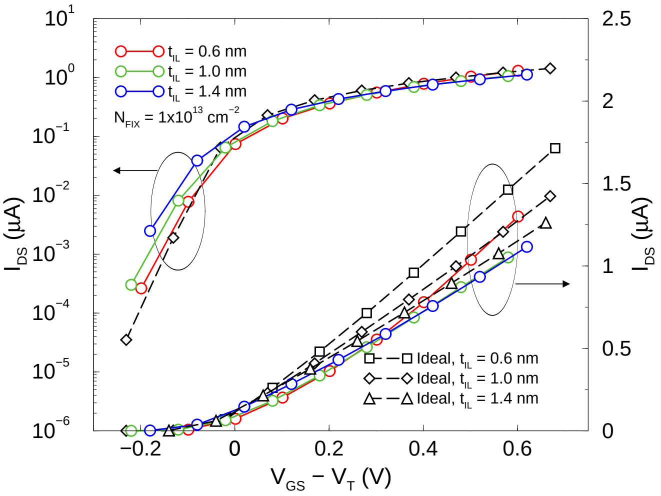

Figure 3.14: *(Solid lines with symbols) Turn-on characteristics for devices with different* t_{IL} *values and a fixed value of* $N_{\text{Fix}} = 1 \times 10^{13}$ *cm*$^{-2}$*. (Black dashed lines with symbols) The reference ballistic cases are reported for comparison. The drain voltage is set to* $V_{\text{DS}} = 5$ *mV.*

induced by the unscreened fixed charges as shown in the left panel of Fig. 3.13, resulting in the V_T shift and enhanced tunneling currents. The right panel of Fig. 3.13 shows the corresponding transmission probabilities as a function of energy. An increasing importance of the tunneling current can be observed for the device with the higher value of N_{Fix}, reflected by a spreading of the current values over almost one order of magnitude in the sub-threshold regime. It is worth highlighting that the transmission probability largely depends on inter-subband scattering due to higher subbands (not shown in the figure), and that this effect is correctly accounted for by the CMS approach. As far as the inverse sub-threshold slope is concerned, a progressive increase from 63 mV/dec to 78 mV/dec is reported, compared to the ideal condition of $SS = 60$ mV/dec of the ballistic device.

RCS also influences the electrical performance at large gate overdrives, where a finite current reduction with increasing N_{Fix} is reported. This is easily explained by the increasing importance of the coupling of transverse modes for large electronic densities: although electronic screening gives rise to a drastic reduction of the spatial fluctuations and carriers experience mostly a thermionic transport, the increased number of conducting modes implies a non-negligible RCS scattering directly proportional to N_{Fix}.

A similar analysis has been carried out for devices characterized by the same $N_{\text{Fix}} = 1 \times 10^{13}$ cm^{-2} value and different thicknesses of the interfacial layer, namely $t_{\text{IL}} = 0.6, 1.0,$ and 1.4 nm. As in the previous case, in the sub-threshold regime, a reduction of performance is reported, namely an increasing I_{OFF} with

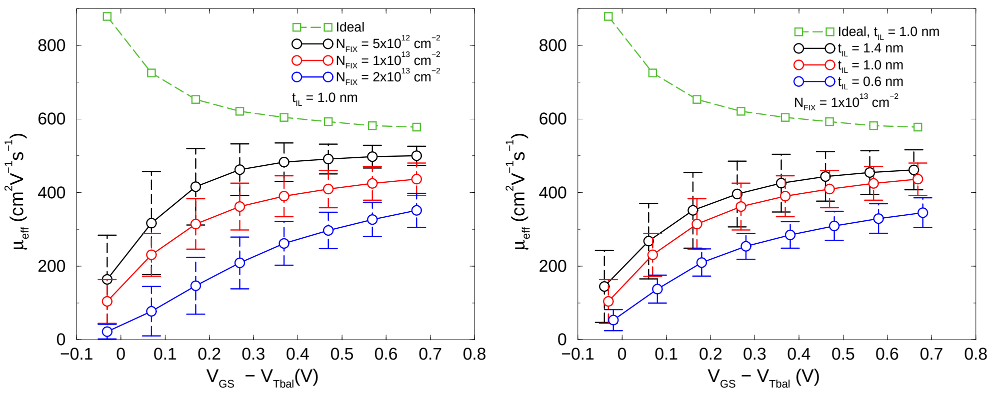

Figure 3.15: *(Left) Effective mobility as a function of* $V_{GS} - V_{T_{bal}}$ *for* $t_{IL} = 1.0$ *nm and for different values of* N_{Fix}*, and (right) for different values of* t_{IL} *and* $N_{Fix} = 1 \times 10^{13}$ *cm*$^{-2}$*. (Green dashed line with symbols) The ballistic mobility curve is reported as reference for the device with* $t_{IL} = 1.0$ *nm.*

t_{IL} due to tunneling currents. As far as t_{IL} is varied, a change in the shape of the potential fluctuations (i.e. frequency and height of the oscillations) is observed, resulting in a higher tunneling probability for the device with a thicker interfacial layer. On the other side, the enhancement of the gate capacitance and the reduction of the distance between fixed charges and carriers in the channel are reflected in a higher impact of RCS on transfer characteristics at large overdrive for the devices with $t_{IL} = 0.6$ and 1.0 nm. A much reduced spreading of threshold voltage values is measured, ranging from $\Delta V_T = 40$ mV for $t_{IL} = 1.4$ nm to $\Delta V_T = 79$ mV for $t_{IL} = 0.6$ nm, with respect to the ideal devices with equal total oxide thickness.

It is important to recall that the analysis was carried out on the transfer characteristics of a single charge distribution for any set of parameters. In such a way, a trend was qualitatively deduced for the comprehension of their impact on transport. A more general result can be derived through the use of a statistical approach, as shown in the following section, focusing on the analysis of the RCS–limited mobility.

3.4.4 RCS–limited mobility

The influence of RCS on the low-field mobility is investigated by using a sample of 10 different realizations of fixed charges due to impurities at the high-κ/SiO$_2$ interface and by a comparison with a reference ballistic device (without fixed charges). The effective mobility extraction method is the same as presented in the previous section.

The mean effective mobilities and their standard deviations are reported in

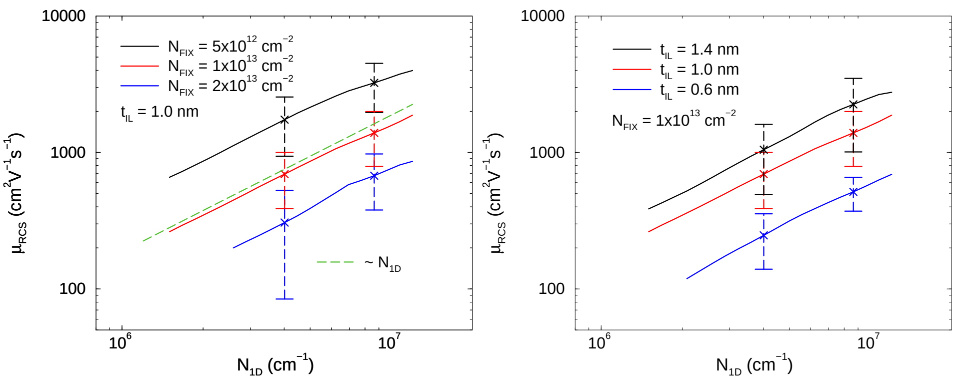

Figure 3.16: *(Left) Mean value of the RCS-limited mobility as a function of the channel electron density, N_{1D}, evaluated over 10 device samples, for different values of $N_{\rm Fix}$ and $t_{\rm IL} = 1.0$ nm, and (Right) for different values of $t_{\rm IL}$ and $N_{\rm Fix} = 1 \times 10^{13}$ cm^{-2}. (Top, Dashed line) The asymptotic behavior $\mu_{\rm RCS} \sim N_{1D}$ is plotted.*

Fig. 3.15 for different values of $N_{\rm Fix}$ and $t_{\rm IL}$. Effective mobilities are plotted as a function of $V_{\rm GS} - V_{T_{\rm bal}}$, where $V_{T_{\rm bal}}$ is the threshold voltage of the reference ballistic device. Due to this choice, the variability of the threshold voltage in each sample is reflected on the spreading of the mobility around the mean value. The largest relative variations are observed in the small overdrive regime where RCS is most effective and the channel electron density is exponentially dependent on the surface potential. According to the discussion on the transfer characteristics, a maximum of the relative spreading is found for the highest $N_{\rm Fix}$ value.

The comparison with the reference ballistic mobility clearly shows the signature of the physical role played by Coulomb interactions with an increase of the effective mobility with the gate overdrive when RCS is suppressed by the screening effect of the charge in the channel. Consequently, larger impact of the RCS on $\mu_{\rm eff}$ is found for the devices with higher density of scattering centers as well as for thinner interface layer thicknesses.

Generally speaking, the effect of RCS on the effective mobility of such short-channel devices strongly depends on $N_{\rm Fix}$ and can be eventually inhibited by optimized technological processes with the production of high quality interfaces.

A non-negligible component of ballistic electrons is still present in the transport regime of the chosen device with gate length of 40 nm, in particular at large overdrive biases. Therefore, in order to analyzed the impact of mere RCS on mobility, the apparent mobility component has been subtracted. Following [29], the decomposition

$$\frac{1}{\mu_{\rm eff}} = \frac{1}{\mu_{\rm bal}(L_{\rm ch})} + \frac{1}{\mu_{\rm RCS}}, \tag{3.6}$$

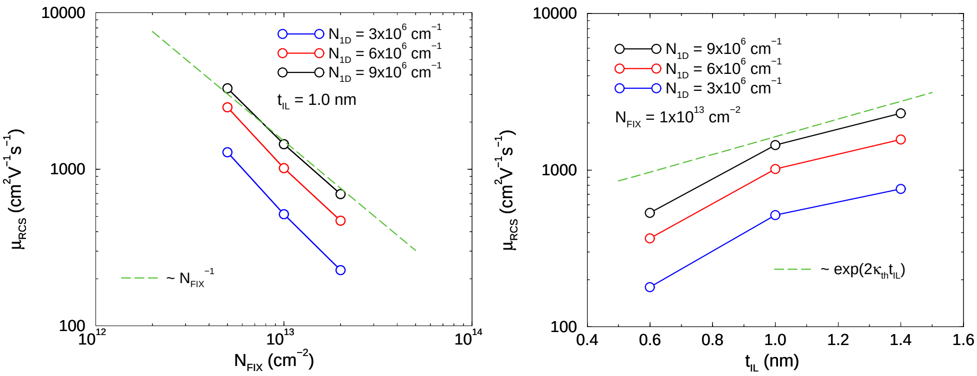

Figure 3.17: *(Left) Mean RCS-limited mobility plotted versus the fixed charges density, N_{Fix}, for different conditions of the channel electron density. (Dashed line) The asymptotic behavior $\mu_{\text{RCS}} \sim 1/N_{\text{Fix}}$ is reported. (Right) Mean RCS-limited mobility plotted versus the interfacial layer thickness, t_{IL}, for different conditions of the channel electron density. (Dashed line) The asymptotic behavior $\mu_{\text{RCS}} \sim \exp(2\kappa_{\text{th}} t_{\text{IL}})$ with $\kappa_{\text{th}} = 0.65$ nm^{-1} is plotted.*

is used, where $\mu_{\text{bal}}(L_{\text{ch}})$ is the ballistic component, which depends on channel length, and μ_{RCS} is the RCS–limited mobility. The above factorization is physically justified since the two scattering mechanisms are independent.

Fig. 3.16 shows the average RCS limited mobility $\left\langle \left(\frac{1}{\mu_{\text{eff}}} - \frac{1}{\mu_{\text{bal}}} \right)^{-1} \right\rangle$ as a function of the linear electron density in the channel. The mean values are extracted on a set of 10 different samples. In order to compare devices at the same inversion charge N_{1D}, an interpolation of the simulated data is performed before averaging. The evaluated standard deviation is reported for the two limit conditions of high and low electron densities.

Both for N_{Fix} and t_{IL} variations a clear power-law dependence of $\mu_{\text{RCS}} \sim N_{1D}^{\alpha}$ is found, where α is close to unity, with a slowly deviation only at high densities. Such a result is coherent with the power law found in experiments on double-gate and ultra-thin-body MOSFETs [22], where the exponent α was observed to increase by reducing the body thickness. Physically, the power-law behavior originates from the screening effects at room temperature as it is given by the increase of the inversion screening length with the 1-D charge density [57].

Finally, in Fig. 3.17 the dependence of the RCS-limited mobility with respect to the two main parameters of the model is analyzed for three different conditions of channel electron density. On one hand, a clear dependence on the inverse of the density of the scattering centers at the high-κ/SiO$_2$ interface is found. The $\mu_{\text{RCS}} \sim 1/N_{\text{Fix}}$ behavior is reported as reference. This result is in agreement with the scattering rate expression given in [57] and can be seen, for the considered scattering parameters and dimensions of the devices, as a validation of the Born approximation used in perturbative models.

On the other hand, a monotonic increase of the RCS–limited mobility with the increasing interfacial layer thickness is found, although a non trivial dependence on t_{IL} is observed. According to experimental results on HfO$_2$/metal gate MOSFETs [23], the empirical law $\mu_{\mathrm{RCS}} \sim \exp(2\kappa_{\mathrm{th}} t_{\mathrm{IL}})$, where κ_{th} is an effective inverse screening length independent of the channel carrier density at room temperature, has been proposed. This reference model is reported for a value of $\kappa_{\mathrm{th}} = 0.65$ nm^{-1}, as proposed in [23]. For sub-nanometric interfacial layers an enhancement of mobility degradation is observed, which is also reported in experimental works (see [58] and works cited therein). The discrepancy with the proposed model for $t_{\mathrm{IL}} < 1.0$ nm can be interpreted as an inefficiency of electron screening when fixed charges are localized at a distance shorter than the effective screening length and by considering their proximity to non-vanishing wave-functions into the oxide.

3.5 Dissipative transport: the impact on device scaling

Although channel lengths in the order of few tens of nanometers are considered, the impact of the electron-phonon interaction can not be totally neglected. As previously observed for the case of CNT-FETs, the inclusion of the fundamental electron-phonon (PH) collision mechanisms have to be considered for accurate estimates of the device performance. For this reason, the analysis with SR and RCS considered so far has been extended here by including the phonon scattering as well. Like in the previous analyses, a linear transport regime is considered, and the attention is focused on the impact of the different scattering mechanisms on transfer characteristics and effective mobility by varying the channel length, L_{ch}. NEGF formalism applied to the electron-phonon interaction in SiNW-FET has been already presented in [59, 60], where the fundamental aspects of a full-quantum treatment of PH scattering were addressed. Here, the aim is to highlight the impact of different scattering sources in short channel devices, showing the important role played by both phase-coherent and phase-breaking mechanisms, using the concepts of effective mobility and apparent mobility introduced in the previous sections.

3.5.1 Transfer characteristic analysis

The turn-on characteristics of devices with different channel lengths are reported in Fig. 3.18 at $V_{\mathrm{DS}} = 5$ mV in both linear and logarithmic scale. Simulations were carried out in presence of only the electron-phonon interaction. Devices have a square cross-section with a lateral width of 5 nm. A gate-all-around architecture

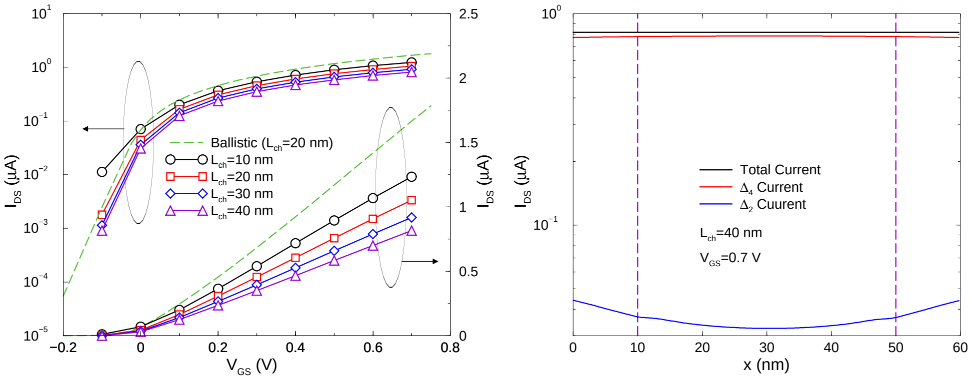

Figure 3.18: *(Left) Turn-on characteristics in presence of PH scattering for different channel lengths.* V_{DS} *is 5 mV. A ballistic reference curve is reported for a device with* $L_{\mathrm{ch}} = 20$ *nm. (Right) Current contributions for the different valleys as a function of position in presence of PH scattering.* $L_{\mathrm{ch}} = 40$ *nm and* $V_{\mathrm{GS}} = 0.7$ *V. The vertical dashed lines indicate the channel-to-contact interface coordinate.*

is used, with a conduction band-edge work-function for the ideal metallic gate. The source and drain contacts have a fixed length of 10 nm, while L_{ch} ranges from 10 to 40 nm. A ballistic reference curve is added, for a device with $L_{\mathrm{ch}} = 20$ nm. The maximum drain current clearly shows a dependence on L_{ch} also for the shorter channel lengths. It is important to note that phonon scattering is considered also in the contact regions, giving an additional contribution to the current reduction. For the shortest device, a non negligible direct source-to-drain tunneling contribution is evident in the sub-threshold regime, which impacts the SS. A shift of the threshold voltage of approximatively 27 mV with respect to the ballistic case is noted for the devices with $L_{\mathrm{ch}} \geq 30$ nm.

As an example, the current contributions of the different valleys, now coupled by inter-valley scattering are reported in Fig. 3.18 (right), along with the total current as a function of the position x, for the device with $L_{\mathrm{ch}} = 40$ nm at $V_{\mathrm{GS}} = 0.7$ V. The Δ_4 and Δ_2 valleys currents are reported separately. The currents associated to each unprimed valley are perfectly equivalent and are represented as a single component of the total current. A current redistribution between the primed and unprimed valleys is observed along the device, while the total current is constant as expected from the continuity equation.

The impact of the additional scattering mechanisms is analyzed. A single configuration of charge distribution and rough interface has been considered for the devices with different gate lengths. For the remote-Coulomb scattering, a defects density of $N_{\mathrm{Fix}} = 1 \times 10^{13}$ cm^{-2} and an interfacial layer of 1.0 nm are considered, while the SR parameters are set to $\Delta_m = 0.2$ nm and $L_m = 1.0$ nm. The

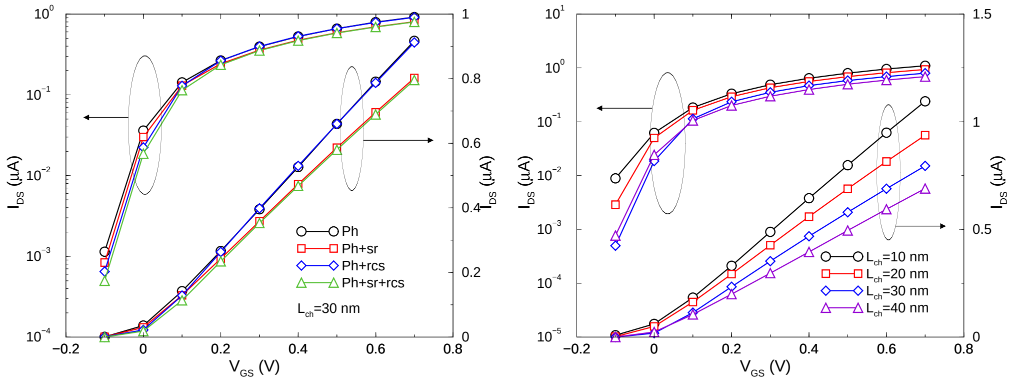

Figure 3.19: *(Left) Turn-on characteristics in presence of different scattering mechanisms for a fixed channel length* $L_{\rm ch} = 30$ *nm.* $V_{\rm DS}$ *is 5 mV. (Right) Turn-on characteristics in presence of all scattering mechanisms considered and different channel lengths.*

transfer characteristics of a device with a channel length of 30 nm are reported in Fig. 3.19. The possible configurations PH+SR, PH+RCS, PH+SR+RCS, have been investigated, showing the different effects as the gate overdrive is varied. In accordance to the previous results both SR and RCS induce a V_T shift, at any gate length.

As far as the remote-Coulomb interaction is concerned, the induced potential fluctuations and the possibility of charge-clustering can causes a variation of the sub-threshold slope. On the contrary, no impact of the surface roughness is reported on SS, even when variations of the channel length are considered, in accordance with the results presented in [14]. SR and RCS scattering have again a complementary behavior as the gate voltage is increased. Due to the screening effect, remote-Coulomb scattering is effective mainly in the sub-threshold region. On the contrary, for the wire size considered in this study, SR has a major impact at high gate overdrives, due to the enhancement of the mode coupling.

The analysis has been extended to the other channel lengths. In Fig. 3.19 (right), the turn-on characteristics of devices with different $L_{\rm ch}$ are reported, computed by incorporating all the scattering mechanisms. The largest impact of RCS in the sub-threshold region is found for the device with $L_{\rm ch} = 30$ nm, reporting a reduction of the current of almost 50% at $V_{\rm GS} = 0$ V with respect the case with PH scattering only. On the contrary, an enhancement of the tunneling current causes a reduction of sub-threshold slope for the device with a channel length of 20 nm. For the devices with $L_{\rm ch} = 10$, and 40 nm a similar reduction of the off-current of almost $15 - 20\%$ is observed, despite the different impact of the DT current. As the gate voltage is increased, a similar impact of SR is found for the different gate lengths, showing a clear signature of an averaging of

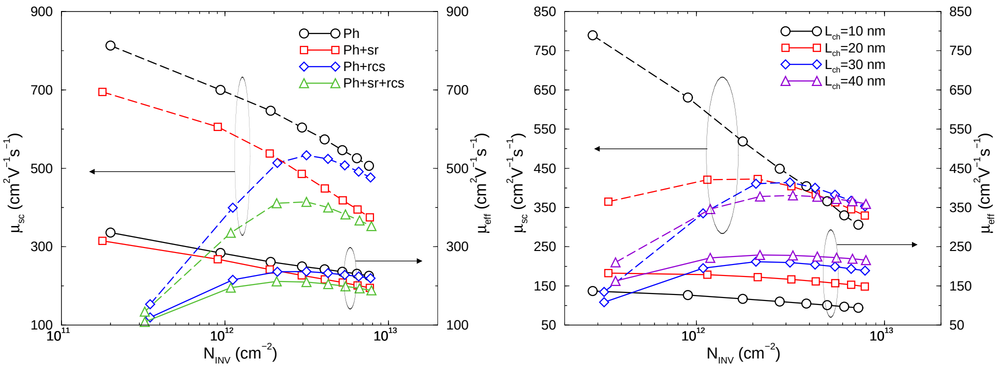

Figure 3.20: *(Left) Effective and scattering–limited mobilities in presence of different scattering mechanisms. The channel length is fixed at $L_{\text{ch}} = 30$ nm. (Right) Effective and scattering–limited mobilities in presence of all considered scattering mechanisms for different channel lengths. Mobility curves are presented as a function of the channel electron density. A normalization with respect to the wire perimeter is supposed.*

the scattering events for $L_{\text{ch}} \geq 20$ nm. A maximum current reduction of 15% with respect the case with PH scattering only is reported for $L_{\text{ch}} = 40$ nm at $V_{\text{GS}} = 0.7$ V.

3.5.2 Short-channel effective mobility

In analogy with the analysis presented on the transfer characteristics, the extraction of the effective mobility clearly shows the signatures of the different scattering mechanisms. In Fig. 3.20 (left) the effective mobility, μ_{eff}, of a single device with channel length $L_{\text{ch}} = 30$ nm is reported as a function of the channel electron density, N_{Inv}. A normalization of the density with respect to the wire perimeter is used. The mobility extractions with different combinations of scattering mechanisms are separately reported. By considering as a reference the mobility given by the electron-phonon interactions, the fundamental aspects observed so far for the SR and RCS mechanisms on mobility are confirmed, and their combination is clearly captured. The RCS affects the effective mobility in the low electron channel density range, while PH scattering and SR are the limiting mechanisms at medium and high channel electron densities, respectively. By applying the mobility decomposition proposed by Shur and presented in the previous section, $\mu_{\text{sc}} = \left({\mu_{\text{eff}}}^{-1} - {\mu_{\text{bal}}}^{-1}\right)^{-1}$, the mobility limited by the channel scattering mechanisms after the factorization of the ballistic or apparent component, is analyzed. A shift of the curves to higher mobility values is observed, being the ballistic component, which depends on the channel length, a limiting factor. No changes are instead observed in the dependence on N_{Inv} for the

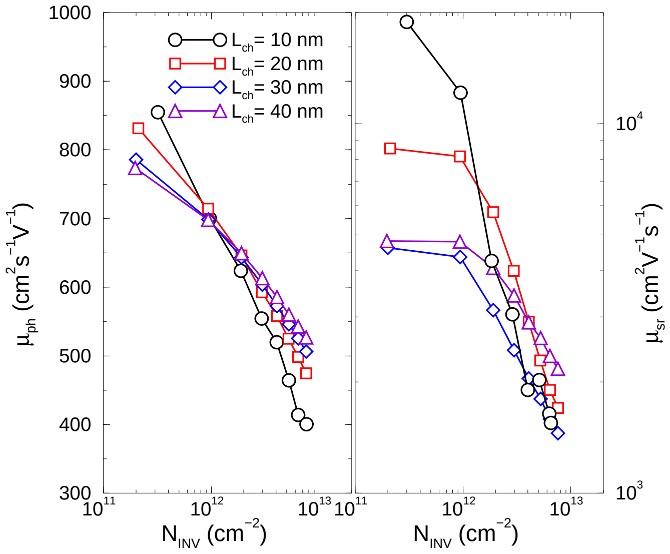

Figure 3.21: *PH–limited (left panel) and SR–limited (right panel) mobility for different values of* L_{ch} *as a function of the channel electron density.*

different scattering mechanisms.

The same analysis has been carried out on channel lengths. Fig. 3.20 (right) shows both the effective mobility and the scattering-limited mobility for different L_{ch} values as a function of N_{Inv}. The effective mobility shows the clear signature of the ballistic component with a progressive lowering of the mobility as the gate length is progressively scaled down, especially at high inversion charges, while the dependence of the curves for low N_{Inv} values is less trivial. Here, a decreased efficiency of the RCS and SR mechanisms makes the mobility degradation of the shortest devices ($L_{ch} = 10$, and 20 nm) less effective at low channel densities.

As the ballistic component of the mobility is removed, a dependence of μ_{sc} on the gate length is in any case observed. This dependence is clear in the low N_{Inv} range. As previously mentioned, results have been collected on devices with a single distribution of fixed charges and of surface roughness for each channel length. Hence, a dependence on the specific occurrence of potential fluctuations has to be taken into account. In any case, it remains clear that SR and RCS show a reduced degradation effect on mobility as the channel length is scaled down. It is more complex and interesting the different dependence on L_{ch} at large gate biases. Here, an inversion of the previous result is observed, with a mobility decreasing with descreasing channel lengths, as shown by the crossing of the different curves in Fig. 3.20. The reasons of the latter behavior are still unclear. One possible explanation is given by the different dependence on N_{Inv} for the different scattering mechanisms.

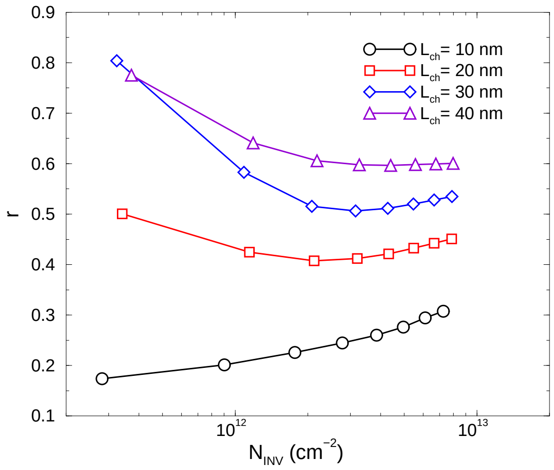

Figure 3.22: *Backscattering coefficient in linear transport regime for different channel lengths. All the considered scattering mechanisms are included.*

In particular, a distinction between phase-breaking and phase-coherent scattering mechanisms needs to be made. To this purpose, the PH–limited mobility, μ_{ph}, and the SR–limited mobility, μ_{SR}, have been isolated, making use of the Matthiessen rule for the mobility decomposition. The results are reported in Fig. 3.21, for different gate lengths over the whole range of N_{Inv}. A reduced variation of mobility for the μ_{ph} for $L_{\text{ch}} \geq 20$ nm is observed if compared with the SR–limited mobilities. In particular, the μ_{SR} mobility curves show a relevant variation with L_{ch} due to the minor impact of SR at low N_{Inv} for the shorter channel lengths and converging to almost the same value at high channel electron densities.

In the case of PH–limited mobility, the faster degradation at high transverse field for the shortest device can be ascribed to an increasing importance of scattering events occurring in the contact regions. In fact, as the gate voltage is increased and the channel potential barrier is lowered, a comparable effect on mobility is expected for both scattering events in the channel and in the source and drain regions. This degradation results obviously to be more evident as L_{ch} is reduced, being the source and drain lengths kept fixed. For the longest device μ_{ph} shows values and a dependence on N_{Inv} similar to these expected for an infinite homogeneous structure [49].

On the other side, a faster degradation of mobility is reported as the channel length is reduced in presence of phase-coherent scattering mechanisms, like SR or RCS. It is important to stress that, in this case, due to the choice of limiting the presence of rough interfaces or charged impurities only in the gated region, no effects on mobility are expected by the source and drain regions. It can be thus concluded that the absence of phase-breaking scattering mechanisms drastically changes the relation between current and charge density in the channel. Furthermore, an enhancement of the effect of carrier degeneracy is observed as we approach the limit of a ballistic transport regime, causing an intrinsic mobility degradation for the short-channel FETs.

As additional analysis of the mobility properties in a quasi-ballistic device has been carried out, based on the determination of the backscattering coefficient in linear transport regime [61], r. Following [62], the backscattering coefficient can be expressed as a function of the effective and the ballistic mobility as

$$r = 1 - \frac{\mu_{\text{eff}}}{\mu_{\text{bal}}} \ , \tag{3.7}$$

where the relation is based on the use of the Matthiessen-like rule for the ballistic component of mobility. The evaluation of the backscattering coefficient as a metric for the impact of scattering on device performance is equivalent to the analysis of the impact of surface roughness presented in Fig. 3.9. The values of r as a function of N_{Inv} are reported in Fig. 3.22, for different gate lengths and in presence of all considered scattering mechanisms. As expected, a progressive decrease of the backscattering coefficient is observed when reducing L_{ch}, which is a clear indication of an increased ballisticity for the short-channel devices. At the same time, a larger relative variation of r is reported at high N_{Inv} for $L_{\text{ch}} = 10$, and 20 nm, confirming the observations drawn on the scattering-limited mobility.

3.6 Summary

In this Chapter, the NEGF formalism within the parabolic effective-mass approximation and the coupled mode space approach has been applied to the analysis of SiNW-FETs. Different scattering mechanisms limiting the performance of ultra-scaled devices have been analyzed, namely: the surface-roughness, the remote-Coulomb, and phonon scattering. SR and RCS are investigated by using a non-perturbative approach, considering devices with specific random realizations of rough Si/SiO$_2$ interfaces and fixed-charge center distributions at the high-κ/SiO$_2$ interface. A general analysis is thus statistically carried out on a set of device samples. The impact of the surface-roughness and remote-Coulomb scattering on the transport properties of SiNW-FETs was analyzed with special attention devoted to the effective mobility. Effective mobility is found to be an important performance metric also in the quasi-ballistic regime, showing findings in accordance with the main semi-classical models. Finally, a global analysis of the interplay of the different scattering mechanisms has been performed showing interesting results on the mobility trend for devices scaled down to 10 nm channel length. Although additional investigations are still envisaged, the realization of a simulation tool capable of giving a full quantum description of transport and of the most relevant scattering mechanisms for post-CMOS Si-based FETs constitutes one of the main results of the presented work.

Bibliography

[1] Y. Cui, and C.M. Lieber, "Functional Nanoscale Electronic Devices Assembled Using Silicon Nanowire Building Blocks", *Science*, vol. 291, no. 5505, pp. 851-853, 2001. [cited at p. 67]

[2] D.D.D. Ma, C.S. Lee, F.C.K. Au, S.Y. Tong and S.T. Lee, "Small-diameter Silicon Nanowire Surfaces," *Science*, vol. 299, no. 5614, pp. 1874-1887, 2003. [cited at p. 67]

[3] K.H. Yeo *et al.*, "Gate-All-Around (GAA) Twin Silicon Nanowire MOSFET (TSNWFET) with 15 nm Length Gate and 4 nm Radius Nanowires", *Tech. Dig. - Int. Elec. Dev. Meet. 2006*, pp. 1-4, 2006. [cited at p. 67]

[4] S.D. Suk, et al., "High performance 5nm radius Twin Silicon Nanowire MOSFET (TSNWFET): fabrication on bulk si wafer, characteristics, and reliability", *Tech. Dig. - Int. Elec. Dev. Meet. 2005*, pp. 717-720, 2005. [cited at p. 67]

[5] R. Wang, H. Liu, R. Huang, J. Zhuge, L. Zhang, D.-W. Kim, X. Zhang, D. Park, and Y. Wang, "Experimental Investigations on Carrier Transport in Si Nanowire Transistors: Ballistic Efficiency and Apparent Mobility", *IEEE Trans. Electron Devices*, vol. 55, pp. 2960-2967, 2008. [cited at p. 67, 68]

[6] C. Dupré *et al.*, "15nm-Diameter 3D Stacked Nanowires with independent Gates operation: ΦFET", in *IEDM Tech. Dig. 2008.* [cited at p. 67]

[7] J. Wang, E. Polizzi, A. Gosh, S. Datta and M. Lundstrum, "Theoretical investigation of surface-roughness scattering in silicon nanowire transistors", *Appl. Phys. Lett.*, vol. 87, p. 043101, 2005 [cited at p. 67, 73, 79]

[8] A. Svizhenko, P.W. Leu, and K. Cho, "Effect of growth orientation and surface roughness on electron transport in silicon nanowires", *Phys Rev. B*, vol. 75, p. 125417, 2007. [cited at p. 67]

[9] K. Natori, "Ballistic metal-oxide-semiconductor field effect transistor" *J. Appl. Phys.*, vol. 76, no. 8, pp. 4879-4890, Jul. 1994. [cited at p. 67]

[10] A. Rahman and M. S. Lundstrom, "A Compact Scattering Model for Nanoscale Double-Gate MOSFET", *IEEE Trans. Electron Devices*, vol. 49, pp. 481-489, 2002. [cited at p. 67, 78]

[11] M. J. Gilbert, R. Akis, and D. K. Ferry, "Phonon-assisted ballistic to diffusive crossover in silicon nanowire transistors" *J. Appl. Phys.*, vol. 98, no. 9, pp. 094303/1-8, Nov. 2005. [cited at p. 67]

[12] B. Yu, L. Wang, Y. Yuan, P. M. Asbeck, and Y. Taur, "Scaling of Nanowire Transistors", *IEEE Trans. Electron Devices*, vol. 55, pp. 2846-2858, 2008. [cited at p. 67]

[13] K. Uchida *et al.*, "Experimental study of carrier transport mechanism in ultrathin-body SOI n- and p-MOSFETs with SOI thickness less than 5 nm", in *IEDM Tech. Dig.*, 2002, pp. 47-50. [cited at p. 67, 82]

[14] M. Luisier, A. Schenk and W. Fichtner, "Atomistic treatment of the interface roughness in Si nanowire transistors with different channel orientations", *Appl. Phys. Lett.*, vol. 90, p. 102103, 2007 [cited at p. 67, 73, 76, 94]

[15] J. Lusakowski *et al.*, "Ballistic and pocket limitations of mobility in nanometer Si metal-oxide semiconductor field-effect transistors", *Appl. Phys. Lett.*, vol. 87, pp. 053507-3, 2005. [cited at p. 67]

[16] A. Martinez *et al.*, "A Self-Consistent Full 3-D Real-Space NEGF Simulator for Studying Nonperturbative Effects in Nano-MOSFETs", *IEEE Trans. Elec. Dev.*, vol. 54, issue 9, pp. 2213-2222, 2007. [cited at p. 67]

[17] E.B. Ramayya, D. Vasileska, S.M. Goodnick and I. Knezevic, "Electron Mobility in Silicon Nanowires", *IEEE Trans. Nano.*, vol 6, no. 1, p. 113, 2007. [cited at p. 67, 69, 73]

[18] S. Jin, T.-W. Tang, M. V. Fischetti, "Simulation of Silicon Nanowire Transistors Using Boltzmann Transport Equation Under Relaxation Time Approximation", *IEEE Trans. Electron Devices*, vol. 55, pp. 727-736, 2008. [cited at p. 67]

[19] R. Kotlyar, M.D. Giles, P. Matagne, B. Obradovic, L. Shifren, M. Stettler, and E. Wang, "Inversion mobility and gate leakage in high-k/metal gate MOSFETs", in *IEDM Tech. Dig.*, 2004, pp. 391-394. [cited at p. 67]

[20] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, "Effective electron mobility in Si inversion layers in metal-oxide-semiconductor systems with a high-κ insulator: The role of remote phonon scattering", *J. Appl. Phys.*, vol. 90, pp. 64587, Nov. 2001. [cited at p. 67]

[21] D. A. Buchanan *et al.*, “80 nm polysilicon gated n-FETs with ultra-thin Al_2O_3 gate dielectric for ULSI applications” *Tech. Dig. - Int. Elec. Dev. Meet. 2000*, pp. 223, 2000. [cited at p. 67]

[22] K. Uchida, J. Koga, and S. Takagi, “Experimental study of carrier transport mechanisms double- and single-gate ultratjin-body MOSFETs –Coulomb scattering, volume inversion, and δT_{SOI}-induced scattering –”, in *IEDM Tech. Dig.*, pp. 805-808, 2003. [cited at p. 67, 91]

[23] M. Cassé et al, “Carrier transport in HfO_2/metal gate MOSFETs: physical insight into critical parameters”, *IEEE Trans. Electron Devices*, vol. 53, pp. 759-768, 2006. [cited at p. 67, 92]

[24] A. Cros *et al.*, “Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling”, in *IEDM Tech. Dig.*, 2006, pp. 663-666. [cited at p. 68]

[25] Y.M. Meziani *et al.*, “Magnetoresistance characterization of nanometer Si metal-oxide-semiconductor transistors”, *J. Appl. Phys.*, vol. 96, no. 10, pp. 5761-5765, Nov. 2004. [cited at p. 68]

[26] K. Romanjek, F. Andrieu, T. Ernst, and G. Ghibaudo, “Characterization of the effective mobility by split C(V) technique in sub 0.1 μm Si and SiGe PMOSFETs”, *Solid-State Electron.*, vol. 49, pp. 721-726, 2005. [cited at p. 68, 76]

[27] W. Chaisantikulwat *et al.*, “Differential magnetoresistance technique for mobility extraction in ultra-short channel FDSOI transistors”, *Solid-State Electron.*, vol. 50, pp. 637-643, 2006. [cited at p. 68]

[28] J. Widiez, T. Poiroux, M. Vinet, M. Mouis, and S. Deleonibus, “Experimental Comparison Between Sub-0.1-m Ultrathin SOI Single- and Double-Gate MOSFETs: Performance and Mobility”, *IEEE Trans. Nanotechnology*, vol. 5, pp. 643-648, 2006. [cited at p. 68]

[29] M.S. Shur, “Low ballistic mobility in submicron HEMT’s,” *IEEE Electron Device Lett.*, vol. 23, pp. 511-513, Sept. 2002. [cited at p. 68, 90]

[30] R. Landauer, “Spatial variation of Currents and Fields due to Localized Scatterers in Metallic Conduction”, *IBM J. Res. Dev.*, vol. 1, p. 233, 1957. [cited at p. 68]

[31] M. Büttiker, “Four-Terminal Phase-Coherent Conductance”, *Phys. Rev. Lett.*, vol. 57, p. 1761, 1986. [cited at p. 68]

[32] M. Zilli, D. Esseni, P. Palestri, and L. Selmi, "On the Apparent Mobility in Nanometric n-MOSFETs", *IEEE Electron Device Lett.*, vol. 28, pp. 1036-1039, Nov. 2007. [cited at p. 68]

[33] K. Huet *et al.*, "Monte Carlo study of apparent mobility reduction in nano-MOSFETs", in *Proc. ESSDERC Conf.*, 2007, pp. 382-385. [cited at p. 68]

[34] K. Nehari, N. Cavassilas, J. L. Autran, M. Bescond, D. Munteanu and M. Lannoo, "Influence of band structure on electron ballistic transport in silicon nanowire MOSFET's: An atomistic study", *Solid-State Electr.*, vol 50, pp. 716-721, 2006. [cited at p. 69, 70, 72]

[35] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. B. Boykin, and G. Klimeck, "Electronic properties of silicon nanowires", *IEEE Trans. Elec. Dev.*, vol. 52, issue 6, pp. 1087-1103, 2005. [cited at p. 69, 70]

[36] Jing Wang, A. Rahman, A. Ghosh, G. Klimeck, M. Lundstrom, "On the validity of the parabolic effective-mass approximation for the I-V calculation of silicon nanowire transistors", *IEEE. Trans. on Electr. Dev.*, vol. 52, p. 1589, 2005. [cited at p. 69, 70]

[37] D. Li *et al.* "Thermal conductivity of individual silicon nanowires", *Appl. Phys. Lett.* , vol. 83, p. 2934, 2003. [cited at p. 69]

[38] I. Ponomareva, D. Srivastava, M. Menon "Thermal conductivity in thin silicon nanowire: phonon confinemnt effect", *Nano Letters* , vol. 7, p. 1155, 2007. [cited at p. 69]

[39] C. T. Huang *et al.* "Er-doped silicon nanowires with 1.54 μm light-emitting and enhanced electrical and field emission properties", *Appl. Phys. Lett.* , vol. 91, p. 093133, 2007. [cited at p. 69]

[40] M. Dovrat, N. Arad, X.-H. Zhang, S.-T. Lee, A. Saar1, "Optical properties of silicon nanowires from cathodoluminescence imaging and time-resolved photoluminescence spectroscopy", *Phys. Rev. B* , vol. 75, p. 205343, 2007. [cited at p. 69]

[41] R. He, P. Yang, "Giant piezoresistance effect in silicon nanowires", *Nature nanotech.* , vol. 53, p. 42, 2006. [cited at p. 69]

[42] A. K. Buin, A. Verma, and M. P. Anantram, "Carrier-phonon interaction in small cross-sectional silicon nanowires". *J. Appl. Phys.*, vol. 104, p. 053716, 2008. [cited at p. 69]

[43] A. K. Buin, A. Verma, and M. P. Anantram, "Significant Enhancement of Hole Mobility in [110] Silicon Nanowires Compared to Electrons and Bulk Silicon". *Nano Lett.*, vol. 8, no. 2, pp. 760–765, 2008. [cited at p. 69]

[44] A. Paul, N. Neophytou, and G. Klimeck, "Orientation dependence of the charge distribution and quantum capacitance in silicon nanowire transistors", in *TECHCON 2008*, 2008, Austin, TX. [cited at p. 70]

[45] http://www.caam.rice.edu/software/ARPACK/, 2002. [cited at p. 72]

[46] S. Jin, M.V. Fischetti, and T.-W. Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: surface-roughness scattering, dielectric screening, and band parabolicity", *J. Appl. Phys.*, vol. 102, p. 083715, Oct. 2007. [cited at p. 72, 73, 75]

[47] S. Jin, M.V. Fischetti, and T.-W. Tang, "Modeling of Surface-roughness scattering in Ultrathin-body SOI MOSFETs", *IEEE Trans. Electr. Dev.*, vol. 54, No 9, pp. 2191–2203, Sept. 2007. [cited at p. 72]

[48] D. Esseni, A. Abramo, L. Selmi, and E. Sangiorgi, "Physically based modeling of low field electron mobility in ultrathin single-and doublegate SOI n-MOSFETs", *IEEE Trans. Electr. Dev.*, vol. 50, no. 12, pp. 2445-2455, Dec. 2003. [cited at p. 72]

[49] M. Lenzi, P. Palestri, E. Gnani, S. Reggiani, A. Gnudi, D. Esseni, L. Selmi, G. Baccarani, "Investigation of the Transport Properties of Silicon Nanowires Using Deterministic and Monte Carlo Approaches to the Solution of the Boltzmann Transport Equation", *IEEE Trans. Electr. Dev.*, vol. 55, no. 8, pp. 2086-2096, Aug. 2003. [cited at p. 72, 97]

[50] S.M. Goodnick, D.K. Ferry, C.W. Wilmsen, Z. Liliental, D. Fathy and O.L. Krivanek, "Surface roughness at the Si(100)-SiO_2 interface", *Phys. Rev. B*, vol. 32, no. 12, pp. 8171–8189, 1985 [cited at p. 74, 79]

[51] Nian Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, "Estimation of the effects of remote charge scattering on electron mobility of n-MOSFETs with ultrathin gate oxides", *IEEE Trans. Electron Devices*, vol. 47, No 2, pp. 440-447, Feb. 2000. [cited at p. 82]

[52] D. Esseni and A. Abramo, "Modeling of electron mobility degradation by remote Coulomb scattering in ultrathin oxide MOSFETs", *IEEE Trans. Electron Devices*, vol. 50, pp. 1665-1674, 2003. [cited at p. 83, 84]

[53] K. Torii, et al., "Effect of interfacial oxide on electron mobility in metal insulator semiconductor field effect transistors with Al_2O_3 gate dielectrics", *Microelectr. Eng.*, vol. 65, pp. 447-453, 2003. [cited at p. 83]

[54] F. Gámiz, J. B. Roldán, J. E. Carceller, and P. Cartujo, "Monte Carlo simulation of remote-Coulomb-scattering-limited mobility in metal-oxide-semiconductor transistors" *Appl. Phys. Lett.*, vol. 82, pp. 3251-3253, 2003. [cited at p. 83]

[55] S. Barraud, L. Thevenod, M. Cassé, O. Bonno, and M. Mouis, "Modeling of Remote Coulomb scattering limited mobility in MOSFET with HfO_2/SiO_2 gate stacks", *Microelec. Eng.*, vol. 84, pp. 2404-2407, 2007. [cited at p. 83]

[56] S. Saito, K. Torii, Y. Shimamoto, O. Tonomura, D. Hisamoto, et al., "Remote Charge Scattering limited mobility in field-effect transistors with $Si0_2$ and $Al_2O_3/Si0_2$ gate stacks", *J. Appl. Phys.*, vol. 98, pp. 113706, 2005. [cited at p. 84]

[57] F. Stern and W. E. Howard, "Properties of semiconductor surface inversion layers in the electric quantum limit", *Phys. Rev.*, vol. 163, pp. 816–835, Oct. 1967. [cited at p. 91]

[58] O. Weber, M. Cassé, L. Thevenod, F. Ducroquet, T. Ernst, and S. Deleonibus, "On the mobility in high-κ/metal gate MOSFETs: Evaluation of the high-κ phonon scattering impact", *Solid-State Elec.*, vol. 50, pp. 626-631, April 2006. [cited at p. 92]

[59] M. Frey, A. Esposito, and A. Schenk "Simulation of intravalley acoustic phonon scattering in silicon nanowires", *European Solid state device Research Conference ESSDERC 2008*, pp. 258-261, 2008 [cited at p. 92]

[60] S. Jin, Y. J. Park, and H. S. Min, "A three-dimensional simulation of quantum transport in silicon nanowire transistor in presence of electron-phonons interactions" *J. Appl. Phys.*, vol. 99, p. 123719, 2006. [cited at p. 92]

[61] Mark Lundstrom, "On the Mobility vs. Drain Current Relation for a Nanoscale MOSFET", *Electr. Dev. Lett.*, vol. 22, pp. 293–295, 2001. [cited at p. 98]

[62] I. Pappas, G. Ghibaudo, C.A. Dimitriadis, and C. Fenouillet-Branger, "Backscattering coefficient and drift-diffusion mobility extraction in short channel MOS devices", *Solide-State Electronics*, vol. 53, pp. 54–56, 2009. [cited at p. 98]

Conclusion

In this thesis, a full-quantum description of transport in the effective-mass approximation was used to simulate post-CMOS nanoscale devices. Both ballistic and dissipative transport regimes have been considered in carbon-nanotube and silicon-nanowire based FETs. Phase-breaking phonon scattering mechanisms and phase-coherent surface-roughness and remote-Coulomb scatterings were considered. Through the use of a Coupled Mode Space approach, a complete tool for the analysis of SiNW-FETs containing all the major limiting transport factors for the latest generations of transistors has been realized.

In the analysis of CNT based devices, improvements to the standard parabolic effective-mass approximation were introduced in order to include, in a computationally advantageous framework, both the non-parabolicity of the energy dispersion relation and the band-to-band tunneling, also in presence of the electron-phonon interaction. The model has been successfully validated via a direct comparison with results obtained with a tight-binding transport model, and applied to analyze the ultimate scaling projections for *n-i-n* and *p-i-n* CNT-FETs.

Although the used models are attractive for their low computational costs and showed a wide range of valid applications, improvements or alternative methods are compulsory when the investigation of the electronic properties of new materials or aggressively scaled and confined channels are considered. The TB method has shown a good trade-off between computational costs and accuracy when applied to carbon-based materials. Its application to really narrow silicon nanowires has given accurate descriptions of the changes in the energy dispersion relation. In both cases, major limitations arise when self-consistent transport analyses are required, imposing constraints on the maximum dimensions of the considered devices. The TB model has in any case to be considered as a reference one, for example in the extraction and calibration of optimal parameters to be used in higher level approximations.

Concerning the description of the electron-phonon interaction, as already mentioned, the effect of confinement of the phonon dispersion relation should be considered. Going beyond the bulk approximation is much more important as the lateral dimensions of silicon nanowire are shrunk. The presence of the coating oxide and other surface alterations should be analyzed also in terms of their effect on silicon and phonon band structure.

Finally, with the growing interest in alternative channel materials and the new possibilities for the fabrication of nanowire heterostructures through the bottom-up approach, such methods, mainly applied to the analysis of silicon-based devices, could find new important applications. Si/Ge or InAs/InP core/shell systems are only two possible examples of heterostructures, attractive mainly for the reported high mobilities. Remaining on the field of carbon-based devices, graphene nanoribbons have recently attracted attention in theoretical and experimental studies. Performances as good as the ones reported for CNTs are expected with the additional advantage of a better compatibility with the standard silicon fabrication process of intrinsically planar devices.

Curriculum Vitae

Stefano Poli was born in Ascoli Piceno, Italy, on October 6, 1981. He received his 1-st and 2-nd level Master Degree (with honors) in Electronic Engineering in 2003 and 2005, respectively. From September 2005 to February 2006 he was visiting student at the IMEC research center, Leuven, Belgium. In 2006, he joined the ARCES research center, University of Bologna, Bologna, starting the European Ph.D. program in Information Technology (EDITH). From September 2007 to September 2008 he was in the position of Early Stage Researcher at the CEA research center, Grenoble, France. He is currently working with the ARCES device modeling and simulation group.

List of Publications

Conference Publications

[1] S. Poli, S. Callegari, R. Rovatti, G. Setti "Post Processing of Data Generated by a Chaotic Pipelined ADC for the Robust Generation of Perfectly Random Bitstreams", IEEE International Symposium on Circuits and Systems (ISCAS 2004), pp. IV-585 IV-588, Vancouver, May 2004.

[2] S. Poli, G. Fiori, S. Reggiani, A. Gnudi, G. Iannacone "Tight-Binding versus Effective-mass modeling of carbon nanotube FETs", 8-th International Conference on Ultimate Integration on Silicon (ULIS), Leuven, March 2007.

[3] S. Poli, S. Reggiani, A. Gnudi, M. Rudan, G. Baccarani "Effective-mass modeling of carbon nanotube FETs", 1-st Italian Workshop on Carbon Nanotubes for Electronics Applications, Bologna, May 2007.

[4] R. Grassi, S. Poli, S. Reggiani, A. Gnudi, M. Rudan, G. Baccarani "Phonon-scattering effects in CNT-FETs with different dimensions and dielectric materials", 37-th European Solid-State Device Research Conference and 33-rd European Solid-State Circuits Conference ESSDERC/ESSCIRC 2007, Munich, September 2007.

[5] R. Grassi, S. Poli, S. Reggiani, E. Gnani, A. Gnudi, G. Baccarani, "Tight-Binding and Effective-mass modeling of armchair carbon nanoribbon FETs", 9-th International Conference on Ultimate Integration on Silicon (ULIS), Udine, March 2008.

[6] C. Kerner, B. Hackens, D. S. Golubovic, S. Poli, S. Faniel, W. Magnus, W. Schoenmaker, V. Bayot, H. Maes, "Quantized logic switch implementation through a superconducting transformer", 10-th International Conference on Ultimate Integration on Silicon (ULIS), Aachen, March 2009.

[7] C. Buran, M. G. Pala, S. Poli, M. Mouis, "Full–3D real–space treatment of surface roughness in double gate MOSFETs", 10-th International Conference on Ultimate Integration on Silicon (ULIS), Aachen, March 2009.

[8] S. Poli and M. G. Pala, "Full quantum investigation of low field mobility in short-channel Silicon Nanowire FETs", submitted to 2009 International Conference on Simulation of Semiconductor Process and Devices (SISPAD).

Journal Publications

[1] S. Poli, S. Reggiani, A. Gnudi, E. Gnani, G. Baccarani, "Computational study of the ultimate scaling limits of the CNT tunnelling devices", IEEE Trans. on Electron Devices, vol. 55, Issue 1, p. 313, 2008.

[2] R. Grassi, S. Poli, S. Reggiani, A. Gnudi, M. Rudan, G. Baccarani, "Phonon-scattering effects in CNT-FETs with different dimensions and dielectric materials", Solid-State Electronics, vol. 52, Issue 9, p. 1329, 2008.

[3] S. Poli, M. G. Pala, T. Poiroux, S. Deleonibus, G. Baccarani, "Size dependence of surface-roughness-limited mobility in silicon-nanowire FETs", IEEE Trans. on Electron Devices, vol. 55, p. 2968, 2008.

[4] C. Kerner, B. Hackens, D. S. Golubovic, S. Poli, S. Faniel, W. Magnus, W. Schoenmaker, V. Bayot, H. Maes, "Control and read-out of current-induced magnetic flux quantization in a superconducting transformer", Supercond. Sci. Technol., vol. 22, p. 025001, Feb. 2009.

[5] R. Grassi, S. Poli, S. Reggiani, E. Gnani, A. Gnudi, G. Baccarani, "Tight-Binding and Effective-mass modeling of armchair carbon nanoribbon FETs", to be published on Solide-State Electronics.

[6] S. Poli, M. G. Pala, T. Poiroux, "Full quantum treatment of Remote-Coulomb scattering in silicon-nanowire FETs", submitted to IEEE Trans. on Electron Devices.